



RT thermophilic archaeobacterium, *Pyrococcus horikoshii* OT3.";

Best Local Similarity 64.9%; Pred. No. 5.8e-82;  
Matches 194; Conservative 39; Mismatches 64; Indels 2; Gaps 1;

```
QY 27 GPDTEAEDVDVWYQTSASAKIDYDKLIVRGSSSIDKELINRIERATGQRPHHRLR 86
Db 23 GGVQVEDEEDVETWYTTKATSIDYDKLIVKGCRCRLDEILARERVRGHKASPLR 82
QY 87 RGIFSHRDMQVLDAYENKRPFLYTGRRGSSAMHGHILPIFTKMLQDVNVLPI 146
Db 83 RGMFAHRDLAIIDRKEGKPFLLYTCRGASSSLHGHILVPIFTKMLQDVNVLPI 142
QY 147 QMTDEKYLKMDLTLQAYDAVENAKDIACGFIDINKTIFESDLQYMGSSGFYKNV 206
Db 143 QMTDEKFLMKMAYDEAKKMARENMKDIISVGFDPKTFIFNNNDY--MCPPYENIVK 200
QY 207 IQKAVTNOVKGIFGFTDSDICGISPAIOAPSFSFQIRDRDIOCLIPCAIDQ 266
Db 201 IMKVVNTNQAAVIFGFPEDCLGKAAPFAVAECFASFPQIFGRNDICLIPCAIDQ 260
QY 267 DPFYRMRDVAAPRIGYKPKALHSTFPFALOGAQTMSASPNSIFLDTAKOIKTKV 325
Db 261 DPFYRMRDVAAPRILKASRPSLIFSTFLPALGQTKMSASBPNTCIFLDTAKOIKTKV 319
```

## RESULT 10

Q976M1 PRELIMINARY; PRT; 406 AA.

```
AC 0976M1:
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DE PUTATIVE TRYPTOPHANYL-TRNA SYNTHETASE.
GN S70169.
OS Sulfolobus tokodaii.
OC Archaea; Crenarchaeota; Sulfolobales; Sulfolobaceae; Sulfolobus.
OX NCBI_Taxid=111955;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JCM 10545 / 7;
RX PubMed=11572479;
RA Kawarabayashi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,
RA Sekine M., Baba S.-I., Ankaï A., Kosugi H., Hosoyama A., Fukui S.,
RA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,
RA Yoshizawa T., Tanaka T., Kudo Y., Yamazaki J., Kishida N., Oguchi A.,
RA Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A.,
RA Oshima T., Kikuchi H.;
RT "Complete genome sequence of an aerobic thermophilic
RT Crenarchaeon, Sulfolobus tokodaii strain7."
RL DNA Res. 8:123-140(2001).
KW EMBL; AP000981; BAB65126.1;
KM Aminoacyl-tRNA synthetase; Hypothetical protein; Complete proteome.
SQ SEQUENCE 406 AA; 47979 MW; 276CCD6506BC94B3 CRC64;
```

Query Match 40.5%; Score 910.5; DB 17; Length 406;  
Best Local Similarity 50.5%; Pred. No. 1.4e-71;  
Matches 189; Conservative 63; Mismatches 99; Indels 23; Gaps 9;

```
QY 35 EDFVDPMTVQTSASAKG-IDYDKLIVRGSSKIDKELINRIERATGQRPHHRLRGI 92
Db 28 QDFVNTVPEV-----KGAVDYDKLIVGFGTOKITELKORIKNLADL-HVMLRRNV 82
QY 93 HRDMNOVLDAVENKRPFLYTGRRGSSAMHGHILPIFTKMLQDVNVLPIQMTDE 152
Db 83 HRDLVLVLDKVDQKGFLLYGRAPSL-GMHIGHILPIFTKMLQDKRNVLLEYITDE 141
QY 153 KYLMK-DLTLQAYDAVENAKDIACGFDINKTIFESDLQYMGSSGFYKNVYKQHV 211
Db 142 KFMNPEYTLDTQROMAVNDILDIIVAGFNDPKYFIPODTEYI--RMVYIAIKIAKL 198
QY 212 TFNVVKGIFGFTDSDICGISPAIOAPSFSNPQIFRRTDIOCLIPCAIDODPYFR 271
Db 199 TFSEVRATFGIDTSNIGITWYPALQIAPT-----MFEKR---KCLIPAGIDODPYFR 248
```

```
QY 272 MTRVAPRIGYKPKALHSTFPFALOGAQTMSADPNSSIFLDTAKOIKTKYKNAFS 331
Db 249 LQRIAESLGYKKAQIHSKRLPPLTGEGKSSSQPETAIIYLDPPKTVARKIMKAFS 308
QY 332 GGRPTIEHROFGNCIDVFSMYLTFPLE-DDDKLEIRKDYSGAMLTGELKALIEV 390
Db 309 GGPPTIELHRRYGGNPDIDVFSQWLYMFEFDDNKIKKIEEDYRSGLTTELKQILIEK 368
QY 391 LQPLIAEQARKK 404
Db 369 LNDPLEHQRKRE 382
```

## RESULT 11

Q97ZX0 PRELIMINARY; PRT; 386 AA.

```
AC 097ZX0:
DT 01-OCT-2001 (TReMBLrel. 18, Created)
DT 01-OCT-2001 (TReMBLrel. 18, Last sequence update)
DE TRYPTOPHANYL-TRNA SYNTHETASE (TRPS) (EC 6.1.1.2).
GN TRPS.
OS Sulfolobus solfataricus.
OC Archaea; Crenarchaeota; Sulfolobales; Sulfolobaceae; Sulfolobus.
OX NCBI_Taxid=2287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35092 / DSM 1617 / P2;
RX MEDLINE=21332296; PubMed=11427726;
RA She Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
RA Aweyer M.J., Chan-Weher C.C.-Y., Clausen I.G., Curtis B.A.,
RA De Moors A., Erasuo G., Fletcher C., Gordon P.M.R.,
RA Helkamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
RA Thi-Noc H.P., Redder P., Schenk M.E., Theinault C., Tolstrup N.,
RA Charlebois R.L., Doellittle W.F., Duguet M., Gaasterland T.,
RA Garrett R.A., Ragan M.A., Sensen C.W., Van der Oost J.;
RT "The complete genome of the crenarchaeon Sulfolobus solfataricus p2."
RL Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
RW EMBL; AE006677; AAK40778.1;
DR InterPro: IPR002305; tRNA-synt_1b.
DR InterPro: IPR002306; tRNA-synt_1b.
DR Pfam: PF00579; tRNA-synt_1b; 1.
DR PRINTS: PRO1039; TRNASYNTHTRP.
KW Aminoacyl-tRNA synthetase; Ligase; Complete proteome.
SQ SEQUENCE 386 AA; 45448 MW; 3F4DA8CB80D219DF CRC64;
```

Query Match 40.4%; Score 907; DB 17; Length 386;  
Best Local Similarity 48.6%; Pred. No. 2.6e-71;  
Matches 191; Conservative 68; Mismatches 110; Indels 24; Gaps 10;

```
QY 31 TEAEDFPVDPMTVQTSASAKG-IDYDKLIVRGSSKIDKELINRIERATGQRPHHRLRGI 89
Db 6 TMDPEFTVPEV-----KGAVDYDKLIVGFGTOKITELKORIKNLADL-HVMLRRNV 59
QY 90 FFSHRDMQVLDAYENKRPFLYTGRRGSSAMHGHILPIFTKMLQDVNVLPIQMT 149
Db 60 FFSHRDLVLVNDYEKSGFLYGRAPSL-GMHIGHILPIFTKMLQDKRNANLLEYIT 118
QY 150 DDEKYLK-DLTLQAYDAVENAKDIACGFDINKTIFESDLQYMGSSGFYKNVYKQ 208
Db 119 DDEKYLK-DLTLQAYDAVENAKDIACGFDINKTIFESDLQYMGSSGFYKNVYKQ 208
QY 209 KHAVTNOVKGIFGFTDSDICGISPAIOAPSFSNPQIFRRTDIOCLIPCAIDODPYFR 268
Db 176 KKLTFSEVRATFGIDASSNIGITWYPALQIAPT-----MFEKR---KCLIPAGIDODPYFR 225
QY 269 YFRMRDVAAPRIGYKPKALHSTFPFALOGAQTMSADPNSSIFLDTAKOIKTKYKN 328
Db 226 YMRLORDIAESLGYKKAQIHSKRLPPLTGEGKSSSQPETAIIYLDVDDPKTVARKIMKY 285
QY 329 AFSGGRDIEHROFGNCIDVFSMYLTFLEDD-DLDEIRKDYSGAMLTGELKAL 387
Db 286 AFSGGPTIELHRRYGGNPDIDVFSQWLYMFEFDDNKIKKIEEDYRSGLTTELKQIL 345
```

RT "The genome sequence of *Drosophila melanogaster*."  
 RL Science 287:2185-2195(2000).  
 DR EMBL: AE003682; AAG22136.1; -.  
 DR FlyBase: FBgn0010803; Aats-trp.  
 DR InterPro: IPR002305; tRNA-synt\_1b.  
 DR InterPro: IPR001412; tRNA-synt\_1.  
 DR InterPro: IPR002306; tRNA-synt\_1.  
 DR Pfam: PF00579; tRNA-synt\_1b; 1.  
 DR PRINTS: PR01039; TRNASYNTHTRP.  
 DR PROSITE: PS00178; AA.TRNA.LIGASE\_I; 1.  
 SQ SEQUENCE 430 AA; 47971 MW; 2338EECC69E979F CRC64;

Query Match 61.4%; Score 1378; DB 5; Length 430;  
 Best Local Similarity 61.4%; Pred. No. 1,3e-112;  
 Matches 262; Conservative 64; Mismatches 85; Indels 16; Gaps 3;

QY 10 DYKADCEPGNAPPTSNCPDA-----TEAE-----EDFDPPTVOTSSAKGIDYD 54  
 DB 3 DTKEVVEGVALTLNGPDAPEVETGDAQAGATAPTDVVDVPMVAVSSNAGVDYD 62  
 QY KLIYFSSSKIDKELINRERATGPRPHFRGIFPSHRMNOVLDAEYENKRPFLYTG 114  
 DB 63 KLIRFSSSKIDDELIAFEKITGRPAHFRGMRFSHRDLITLIREGKFFLYTG 122  
 QY 115 RGPSSAMHVGHLIPFTTKLVQDVFNVLVQMTDEKYLMDLTDQAYGDAVENAKD 174  
 DB 123 RGPSSGLHVGHVLPVMTKWLQETFDVPLVQLTDEKTLMDLTKVEDAIKLGRENKD 182  
 QY 175 ILAGGPDINKFTISDLDYMGSSGFYKNVYKIQKHTFNQYKGIFFGTTSDCIGKISFP 234  
 DB 183 IVAIGFVNKFTIFNNLEFVGKCPAMYONIRIQKCVFNFVQKGIFFGSDIIGKIGFP 242  
 QY 235 AIOAPFSNFPDIQIRTDIOCLIPCAIDODPYFRMTFRVAVRIGYPRKALLHSFPF 294  
 DB 243 AAOAPALISSTFFPIFGNR-KVHCLIPCAIDODPYFRMTFRVAVRIGYPRKALLHSFPF 301  
 QY 295 ALQAGQTKMSASDPNSSIFLTDIAKQIKTKYKNKHAFGSGRDTIEHRQFGNCDVDSFM 354  
 DB 302 ALQAGTKMSASDPNSSANVYLTDPKQIKNKINKYAFSGRVSVEHRKLGVPEDVSYQ 361  
 QY 355 YLTFLEDDDKLEQIKRDYTGAMLTGELKALLIEVLQPLAEHQARKKETDELVEKFM 414  
 DB 362 LKTFLEDDDALEVRVAYSKGMLEGEIKLAVETLPIVEHQARKKLTDELVEKYF 421  
 QY 415 TPKLISF 421  
 DB 422 ELRLPKF 428

RESULT 8  
 Q9SR15 PRELIMINARY; PRT; 402 AA.  
 AC Q9SR15;  
 DT 01-MAY-2000 (TReMBLrel. 13, Created)  
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)  
 DT 01-DEC-2001 (TReMBLrel. 13, Last annotation update)  
 DE PUTATIVE TRYPTOPHANYL-TRNA SYNTHETASE.  
 GN F7018.7.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-CV. COLUMBIA;  
 RA Lin X., Kaul S., Town C.D., Beutlo M., Creasy T.H., Haas B.,  
 RA Rensing C.M., Koo H., Fujii C.Y., Uterback T.R., Barnstead M.E.,  
 RA Bowman C.L., White O., Niernan W.C., Fraser C.M.;  
 RA "Arabidopsis thaliana chromosome III BAC F7018 genomic sequence."  
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AC011437; AAF04890.1; -.  
 DR InterPro: IPR002305; tRNA-synt\_1b.

DR InterPro: IPR001412; tRNA-synt\_1.  
 DR InterPro: IPR002306; tRNA-synt\_trp.  
 DR Pfam: PF00579; tRNA-synt\_1b; 1.  
 DR PRINTS: PR01039; TRNASYNTHTRP.  
 DR PROSITE: PS00178; AA.TRNA.LIGASE\_I; 1.  
 KW Aminoacyl-tRNA synthetase.  
 SQ SEQUENCE 402 AA; 45754 MW; B9BEA75EE5D6CD15 CRC64;

Query Match 57.9%; Score 1301; DB 10; Length 402;  
 Best Local Similarity 60.3%; Pred. No. 6.8e-106;  
 Matches 240; Conservative 66; Mismatches 88; Indels 4; Gaps 2;

QY 29 DATEAE--EDFVDPPTVOTSSAKGIDYDKLIYFSSSKIDKELINRERATGPRPHFR 86  
 DB 7 DERAESESEGVVNWWEVSARKGKIDYDKLIDKRGCCQLDSESLDRVQRLSRQPHVELR 66  
 QY 87 RGIFFSHRDMNOVLDAENKRPFLYITGRGSSSEAMHVGHLIPFTTKLVQDVFNVLVY 146  
 DB 67 RSVFARHDFNEILDAVERGDKFYLYTGRGSSSALHGLIPFTTKYLOEAKFVPLVI 126  
 QY 147 QMTDEKYLMDLTDQAYGDAVENAKDIACGFDINTFTFSDLDYMGSSGFYKNVYK 206  
 DB 127 QLTDEKSIKNSLVSESQRLARENARDIACGFDVTKTFTFSDPDVYG--GAFYKNVYK 184  
 QY 207 IQKHTFNQYKGIFFGTTSDCIGKISFPALQAPFSNSFPQIRFRDTIOCLIPCAIDQ 266  
 DB 185 VGKCVTLNKAMGIFGSESDPIALSPFPVQAVSPSSFPHLPRGKNLRLIPCAIDQ 244  
 QY 267 DPTFRMTRDVAPRIGYPRKALLHSFPALQAGQTKMSASDPNSSIFLTDIAKQIKTKVN 326  
 DB 245 DPTFRMTRDVAPRIGYSPALISSTFFPALQGENKMSASDPNSAIYVTDIAKQIKTKVN 304  
 QY 327 KHAFGSGRDTIEHRQFGNCDVDSFMVLFPEDDDKLEQIKRDYTGAMLTGELKKA 386  
 DB 305 RYAFSGGSDSTEKRELGANLEVDIPKYLSFFLEDDSELHINKETGEGMLTGEVKKR 364  
 QY 387 LIEVLQPLAEHQARKKETDELVEKFMTPRKLSFDPQ 424  
 DB 365 LIEVLTEIVERHRRARAAVTDEMVAFAVAPRLPSMFE 402

RESULT 9  
 Q9U1R2 PRELIMINARY; PRT; 324 AA.  
 AC Q9U1R2;  
 DT 01-MAY-2000 (TReMBLrel. 13, Created)  
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)  
 DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)  
 DE Y80D3A.1 PROTEIN.  
 GN Y80D3A.1.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Harris B.R.;  
 RA Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=99069613; Pubmed=9851916;  
 RA none;  
 RT "Genome sequence of the nematode *C. elegans*: A platform for  
 RT investigating biology."  
 RL Science 282:2012-2018(1998).  
 DR EMBL: AL132853; CAB60439.1; -.  
 DR InterPro: IPR002305; tRNA-synt\_1b.  
 DR InterPro: IPR002306; tRNA-synt\_trp.  
 DR Pfam: PF00579; tRNA-synt\_1b; 1.  
 DR PRINTS: PR01039; TRNASYNTHTRP.  
 SQ SEQUENCE 324 AA; 36289 MW; 6E687E6D420ECC12 CRC64;

Query Match 45.7%; Score 1027; DB 5; Length 324;



QY	115	RGSSSEAMHVGHLIDPEFTFKKLDGVNPLVIOGTMDEKYLMLKDLTDQAYGDAVENAKD	174
QY	115	RGSSSEAMHVGHLIDPEFTFKKLDGVNPLVIOGTMDEKYLMLKDLTDQAYGDAVENAKD	174
Db	123	RGSSGSLHVGHLVPELMTKLLQETFDVPLVIOGTMDEKYLMLKDLTDQAYGDAVENAKD	182
QY	175	IIACGFDFINTKFFIEDDDLYOMGMSGSEGVYNNVKIQKHNTFNOVKQIGFQFTSDICGKISFP	234
Db	183	IVALGFDVNTKFFIFNNLFEVQKCCPAMQONIIIRIOKCTFQNVYKQIFBFGSDIITGKIGFP	242
QY	235	AIQAAFSFNSNFPQIIFRDRITDQCLIPCADIDQDPEFRMTQDVAPRIGYPRPALLHSTFFP	294
Db	243	AAQAAPAFISSTFFPEFIFGNR-KVHCLIPCAIDQDPEFRMTQDVAPRIGYPRPALLHSTFFP	300
QY	295	ALQAGQATKMSADPNSSITFLDTFAKQIKTKYKNKAFFSGNGRPTIEHQFGONGCDVDVSEM	354
Db	302	ALQAGKATKMSADPNSSITFLDTFAKQIKTKYKNKAFFSGNGRPTIEHQFGONGCDVDVSEM	360
QY	355	YLTFLEEDDDLEQIRKQDYTSGAMLTGELKALLLEVLOPLAEHQARRKETDEIVEFPM	414
Db	362	LKTFLEEDDALLEFRVAVSKGEMLTGELKALLLEVLOPLAEHQARRKETDEIVEFPM	421
QY	415	TPRKLSF 421	
Db	422	ELRPLKF 428	

RESULT 6

Q904Y0 PRELIMINARY; PRT; 420 AA.

Q904Y0

AC 01-MAY-2000 (TREMBLrel. 13, Created)

RT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)

DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)

DE TRIPTOPHANYL-TRNA SYNTHETASE (FRAGMENT).

GN AAT5-TRP OR CG9735.

OS Drosophila melanogaster (Fruit fly).

CC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC Pterygota; Neoptera; Endopterygota; Diptera; Braachyera; Muscomorpha;

OX Empidoidea; Drosophilidae; Drosophila.

NCBI\_TaxID=7227;

RM SEQUENCE FROM N.A.

RP MEDLINE=99250164; PubMed=10233165;

RA Seshalah P., Andrew D.J.;

RT "WMS-65D: A tryptophanyl-tRNA synthetase expressed to high levels in the developing prothoracic salivary gland."

RL Mol. Biol. Cell 10:1595-1608(1999).

DR EMBL; AF125157; AAF20167.1; -.

DR FLYbase; FBgn0010803; Aats-trp.

DR InterPro; IPR002305; tRNA-synt\_1b.

DR InterPro; IPR001412; tRNA-synt\_1.

DR InterPro; IPR002306; tRNA-synt\_trp.

DR Pfam; PF00579; tRNA-synt\_1b; 1.

DR PRINTS; PR01039; TRNASYNTHTRP.

DR PROSITE; PS00178; AA\_tRNA\_LIGASE\_I; 1.

DR KW Aminoacyl-tRNA synthetase.

FT NON TER 1

SO SEQUENCE 420 AA; 46975 MW; 4AF7088AF426AE6 CRC64;

Query Match 61.4%; Score 1379.5; DB 5; Length 420;

Best Local Similarity 62.7%; Pred. No. 9.1e-113;

Matches 257; Conservative % 62; Mismatches 90; Indels 1; Gaps

QY 12 KADPRGNRPAFTSNIGRPAATEAEEDVDPWQWYSSAKGIDYKLLYRFGSSKIDKRLIN 71

Db 10 KDAEVRVETGTDQAQESQATPTEDEVVDPMVAVSSNAGVDYDGLIRFSSSKIDEELIA 69

QY 72 RIERATGQRPHNHFLEKRGIFESHRDMQNVLDAYEAKKPFYLYTGRGSSSEAMHVGHLIPT 131

Db 70 RPEKTKGRANHFLEKRGIFESHRDMQNVLDAYEAKKPFYLYTGRGSSSGLHVGHLVPLI 129

QY 132 FTKKLDGVNPLVIOGTMDEKYLMLKDLTDQAYGDAVENAKDIIACGFDFINTKFFISDL 191

5

Db	130	MTNQLGEFFDVLPVLIQQLDDEKTLMLKDLKADAKIKLGRENAKOIVALGPDVNNKTFIFNNL	189
Oy	192	DIMGSSCFYKNVYKIQKHTFNOVYKIGFGETSDCIKISFPALQAAPSFNSPQIFR	251
Db	190	EFVGGCPAMYNIIIRIOCKTFPNVYKIGFGCGSDIIGIKGFPAAQAAPISSTFPPIFG	249
Oy	252	DRTDIOCLIPCALIDDDPYFRMTROVAPRIGCPKRALHSTFFPALQQAOTKMSASPNSS	311
Db	250	NR-KVHCILIPCALIDDDPYFRMTROVAPRIGCPKRALHSTFFPALQQAOTKMSASPNSSA	308
Oy	312	IFLDTIAOIKTKYKNAHFAFSGSGRPTIEEHRQFGNCNVDSFVWLFFLEDDDKLEQIR	371
Db	309	YLDTTPQOIKNKIKYKAFSGSGRTVEEHRKLGVPENDVSYOLKLFKFLDDAKLEVRV	368
Oy	372	DYTSAGMLTGLKALIEVLIOPLIAEHQARREYTDIEVKEFMTPKLSF	421
Db	369	AVSKGEMLTGEIKRLAVETLPRIVEHQARKLITDEVLDRKLEPRLKE	418
RESULT 7			
ID	09VHG2	PRELIMINARY:	PRT: 430 AA.
AC	09VHG2:		
DT	01-MAY-2000 (TREMBLrel. 13, Created)		
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)		
DT	01-JUN-2001 (TREMBLrel. 17, Last annotation update)		
DE	AATS-TRP PROTEIN.		
GN	AATS-TRP OR CG9735.		
OS	Drosophila melanogaster (fruit fly).		
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;		
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;		
OC	Ephydroidea; Drosophilidae; Drosophila.		
OX	NCBI_TaxID=7227;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-BERKELEY.		
RX	MEDLINE-20196006; PubMed=10731132;		
RA	Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,		
RA	Amend M.D., Celniker S.E., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,		
RA	Georgieva R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,		
RA	Sutton G.G., Wortman J.R., Vandell M.D., Zhang Q., Chen L.X.,		
RA	Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,		
RA	Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Mikos G.L.G.,		
RA	Abriil J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,		
RA	Ballew R.M., Basu A., Baxendale J., Bayraktarglu L., Beasley E.M.,		
RA	Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,		
RA	Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,		
RA	Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,		
RA	Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,		
RA	de Pablos K., Delcher A., Deng Z., Duan A.D., Dew I., Dietz S.M.,		
RA	Dodson K., Doup L.E., Downes M., Mayano-Rocha S., Dunkov B.C., Dunn P.,		
RA	Dubin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,		
RA	Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,		
RA	Glodex A., Gong F., Gottrell J.H., Gu Z., Guan P., Harris M.,		
RA	Harris N.L., Haveway D., Heiman T.J., Hernandez J.R., Houck J.,		
RA	Hoslin D., Houston K.A., Howland T.J., Wei M.-H., Idegami C.,		
RA	Jajalia M., Kalush F., Karen G.H., Ke Z., Kennison J.A., Ketchum K.A.,		
RA	Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kuip D., Lai Z.,		
RA	Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,		
RA	Liu X., Maitel B., McIntosh T.C., McLeod M.P., McPherson D.,		
RA	Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,		
RA	Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,		
RA	Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacible J.M.,		
RA	Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,		
RA	Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,		
RA	Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,		
RA	Spieler E., Spradling A.C., Stemple M., Strong R., Sun E.,		
RA	Svitsky S.R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,		
RA	Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weisenbach J.,		
RA	Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,		
RA	Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,		
RA	Zheng X.H., Zhong F.-N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,		
RA	Gibbs R.A., Myers E.W., Rhoad G.M., Venter J.C.;		

Db	130	MTNQLGEFFDVLPVLIQTLDEDKTLMKDLKAVDAKIKLGRENKAKOLVALGPDVNNKTFIFNNL	189
Oy	192	DIMGSSCFYKNVYKIQKHTVFNQVKGICFGETSDCIKISFPALQAAPSFNSPQIFR	251
Db	190	EFVGGCPAMYNIIIRIOKCFYFNQVKGICFGGSDSDIIKIGIFPAQAAPSSSTFPPIFG	249
Oy	252	DRTDIOCLIPCALIDDDPYFRMTROVAPRIGCPKRALHSTFFPALQQAOTKMSASPNSS	311
Db	250	NR-KVHCILIPCALIDDDPYFRMTROVAPRIGCPKRALHSTFFPALQQAOTKMSASPNSSA	308
Oy	312	IFLDTIAKOITKTKYKNAHFAFSGSGRPTIEEHRQFGNCNVDSFVLFLEDDDDLEQIR	371
Db	309	YLDTTPQIOIKNKIKYKAFSGSGRTVEEHRKLGVPENDVSYOLKLFKFLDDAKLEVRV	368
Oy	372	DYTSAGMLTGLKALIEVLIOPLIAEHQARREVTDEIVAEFMTPKLSF	421
Db	369	AVSKGEMLTGEIKRLAVETLPRIVEHQARKLITDEVLDRKLEPRLKE	418
RESULT 7			
ID	O9VHG2	PRELIMINARY:	PRT: 430 AA.
AC	O9VHG2:		
DT	01-MAY-2000 (TREMBLrel. 13, Created)		
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)		
DT	01-JUN-2001 (TREMBLrel. 17, Last annotation update)		
DE	AATS-TRP PROTEIN.		
GN	AATS-TRP OR CG9735.		
OS	Drosophila melanogaster (fruit fly).		
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;		
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;		
OC	Ephydroidea; Drosophilidae; Drosophila.		
OX	NCBI_TaxID=7227;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-BERKELEY.		
RX	MEDLINE-20196006; PubMed=10731132;		
RA	Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,		
RA	Amend M.D., Celniker S.E., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,		
RA	Georgieva R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,		
RA	Sutton G.G., Wortman J.R., Vandell M.D., Zhang Q., Chen L.X.,		
RA	Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,		
RA	Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Mikos G.L.G.,		
RA	Abriil J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,		
RA	Ballew R.M., Basu A., Baxendale J., Bayraktarglu L., Beasley E.M.,		
RA	Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,		
RA	Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,		
RA	Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,		
RA	Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,		
RA	de Pablos K., Delcher A., Deng Z., Duan A.D., Dew I., Dietz S.M.,		
RA	Doeson K., Doup L.E., Downes M., Mayano-Rocha S., Dunkov B.C., Dunn P.,		
RA	Dubin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,		
RA	Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,		
RA	Glodek A., Gong F., Gottrell J.H., Gu Z., Guan P., Harris M.,		
RA	Harris N.L., Haveway D., Heiman T.J., Hernandez J.R., Houck J.,		
RA	Hoslin D., Houston K.A., Howland T.J., Wei M.-H., Ideyang C.,		
RA	Jajalia M., Kalush F., Karen G.H., Ke Z., Kennison J.A., Ketchum K.A.,		
RA	Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kuip D., Lai Z.,		
RA	Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,		
RA	Liu X., Maitel B., McIntosh T.C., McLeod M.P., McPherson D.,		
RA	Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,		
RA	Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,		
RA	Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacible J.M.,		
RA	Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,		
RA	Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,		
RA	Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,		
RA	Spieler E., Spradling A.C., Stemple M., Strong R., Sun E.,		
RA	Svitsky S.R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,		
RA	Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weisenbach J.,		
RA	Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,		
RA	Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,		
RA	Zheng X.H., Zhong F.-N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,		
RA	Gibbs R.A., Myers E.W., Zbign G.M., Venter J.C.;		

RA Sasaki H., Sato K., Schoenbach C., Seta T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilmshurst L.,  
 RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kotsuki S.,  
 RA Hayashizaki Y.,  
 RT "Functional annotation of a full-length mouse cDNA collection."  
 RL Nature 409:685-690(2001).  
 DR EMBL: AK007754; BAB5235.1; -  
 DR MGD: MGI:104630; Mars.  
 DR InterPro: IPR002305; tRNA-synt\_1b.  
 DR InterPro: IPR001412; tRNA-synt\_1.  
 DR InterPro: IPR002306; tRNA-synt\_trp.  
 DR Pfam: PF00579; tRNA-synt\_1b; 1.  
 DR PRINTS: PR01039; tRNA-synt-trp.  
 DR PROSITE: PS00178; AA\_TRNA\_LIGASE\_I; 1.  
 SQ SEQUENCE 329 AA; 37613 MW; 701E702DC244CA2A CRC64;

Query Match 72.5%; Score 1629; DB 11; Length 329;  
 Best Local Similarity 92.1%; Pred. No. 7.7e-135;  
 Matches 303; Conservative 17; Mismatches 9; Indels 0; Gaps 0;

QY 96 MNOVDAYENKRPFLYGRGPSSAMHVLPIFTKMLQDVFNPLVLOMTDDEKYL 155  
 DB 1 MNOIIDAYENKRPFLYGRGPSSAMHVLPIFTKMLQDVFNPLVLOMTDDEKYL 60  
 QY 156 WKDLTLDAYDAVENAKDIIACGFDINKTEIFSDLDYMGSSGFRKVVVIOKHVTFNQ 215  
 DB 61 WKDLTLDAYDAVENAKDIIACGFDINKTEIFSDLDYMGSSGFRKVVVIOKHVTFNQ 120  
 QY 216 VKGIFGFSDSCIGKISFPAIQAPSFNSFPQIFRDTIQCLIPCAIDDPYFRMTRD 275  
 DB 121 VKGIFGFSDSCIGKISFPAIQAPSFNSFPQIFRDTIQCLIPCAIDDPYFRMTRD 180  
 QY 276 VAPRIGYKPPALHSTFEPALOGAOTKMSADPNSTIFLTPTAKOIKKVKHAFSGGRD 335  
 DB 181 VAPRIGYKPPALHSTFEPALOGAOTKMSADPNSTIFLTPTAKOIKKVKHAFSGGRD 240  
 QY 336 TIEEHROFGNCDVDVSEMYLTFLEDDDKLEQIRKDYTSGAMLTGELKALIEVLOPLI 395  
 DB 241 TIEEHROFGNCDVDVSEMYLTFLEDDDKLEQIRKDYTSGAMLTGELKALIEVLOPLI 300  
 QY 396 AEHQARRKEVTDVYKEMTPRKISFDFQ 424  
 DB 301 AEHQARRKEVTDVYKEMTPRKISFDFQ 329

RESULT 4  
 070184 PRELIMINARY; PRT; 305 AA.  
 ID 070184  
 AC 070184  
 DT 01-AUG-1998 (Tremblrel. 07, Created)  
 DT 01-AUG-1998 (Tremblrel. 07, Last sequence update)  
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)  
 DE TRYPHOPHAN-TRNA SYNTHETASE (FRAGMENT).  
 OS Cavia porcellus (Guinea pig).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Hystriognathii; Caviidae; Cavia.  
 OX NCBI\_Taxid=10141;  
 RN [1]  
 RE SEQUENCE FROM N.A.  
 RC STRAIN=HARTLEY; TISSUE=SPLEEN;  
 RA Yang D., Goto R., Watanabe N., Kobayashi Y.;  
 RT "Identification and Cloning of Genes Whose Expressions are Elevated  
 RT during DMCB-Induced Guinea Pig Skin Delayed-type Hypersensitivity  
 RT Reaction."  
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AB012222; BAA25288.1; -  
 DR InterPro: IPR002305; tRNA-synt\_1b.  
 DR InterPro: IPR002306; tRNA-synt\_trp.  
 DR Pfam: PF00579; tRNA-synt\_1b; 1.  
 DR PRINTS: PR01039; tRNA-synt-trp.  
 DR Aminoacyl-tRNA synthetase.  
 KW NON\_TER  
 FT SEQUENCE 305 AA; 34838 MW; 0937164333780EB7 CRC64;

Query Match 68.4%; Score 1537; DB 11; Length 305;  
 Best Local Similarity 94.4%; Pred. No. 8.4e-127;  
 Matches 288; Conservative 11; Mismatches 6; Indels 0; Gaps 0;

QY 120 EAMHGVHLPIFTKMLQDVFNPLVLOMTDDEKYLKDLTLDAYDAVENAKDIIACG 179  
 DB 1 EAMHGVHLPIFTKMLQDVFNPLVLOMTDDEKYLKDLTLDAYDAVENAKDIIACG 60  
 QY 180 FDIKKTFFSDLDYMGSSGFRKVVVIOKHVTFNQKGFIFGFSDSCIGKISFPAIQAA 239  
 DB 61 FDIKKTFFSDLDYMGSSGFRKVVVIOKHVTFNQKGFIFGFSDSCIGKISFPAIQAA 120  
 QY 240 PSFNSFPQIFRDTIQCLIPCAIDDPYFRMTRDVAAPRIGYKPPALHSTFEPALOGA 299  
 DB 121 PSFNSFPQIFRDTIQCLIPCAIDDPYFRMTRDVAAPRIGYKPPALHSTFEPALOGA 180  
 QY 300 QTKMSADPNSTIFLTPTAKOIKKVKHAFSGGRDTIEEHROFGNCDVDVSEMYLTF 359  
 DB 181 QTKMSADPNSTIFLTPTAKOIKKVKHAFSGGRDTIEEHROFGNCDVDVSEMYLTF 240  
 QY 360 LEDDDKLEQIRKDYTSGAMLTGELKALIEVLOPLIAEHQARRKEVTDVYKEMTPRK 419  
 DB 241 LEDDDKLEQIRKDYTSGAMLTGELKALIEVLOPLIAEHQARRKEVTDVYKEMTPRK 300  
 QY 420 SFDFQ 424  
 DB 301 SFDFQ 305

RESULT 5  
 0904Y1 PRELIMINARY; PRT; 430 AA.  
 ID 0904Y1  
 AC 0904Y1  
 DT 01-MAY-2000 (Tremblrel. 13, Created)  
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)  
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)  
 DE TRYPHOPHAN-TRNA SYNTHETASE.  
 GN AATS-TRP OR CG9735.  
 OS Drosophila melanogaster (Fruit fly).  
 CC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 CC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 CC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_Taxid=7227;  
 RN [1]  
 RE SEQUENCE FROM N.A.  
 RC MEDLINE=99250164; PubMed=10233165;  
 RA Seshalah P., Andrew D.J.;  
 RT "WRS-85D: A tryptophanyl-tRNA synthetase expressed to high levels in  
 RT the developing Drosophila salivary gland."  
 RL Mol. Biol. Cell 10:1595-1608(1999).  
 DR EMBL: AF125156; AAR20166.1; -  
 DR FlyBase: FBgn0010803; Aats-trp.  
 DR InterPro: IPR002305; tRNA-synt\_1b.  
 DR InterPro: IPR001412; tRNA-synt\_1.  
 DR InterPro: IPR002306; tRNA-synt\_trp.  
 DR Pfam: PF00579; tRNA-synt\_1b; 1.  
 DR PRINTS: PR01039; tRNA-synt-trp.  
 DR PROSITE: PS00178; AA\_TRNA\_LIGASE\_I; 1.  
 KW Aminoacyl-tRNA synthetase.  
 SQ SEQUENCE 430 AA; 47985 MW; 2E3F9E9BC1E9979A CRC64;

Query Match 61.5%; Score 1382; DB 5; Length 430;  
 Best Local Similarity 61.6%; Pred. No. 5.7e-113;  
 Matches 263; Conservative 63; Mismatches 85; Indels 16; Gaps 3;  
 QY 10 DYKADCCPGNAPATNSNGPDA-----TEAE-----EDFYDPTVOTSSAKGIDYD 54  
 DB 3 DTKETIVEGVALTLNKGPDPEVETGTDAQAQAGATAPEDVDVDPNNVASSNDAGDYD 62  
 QY 55 KLIVRFGSSKIDKELINRIERATGQRPHFLRGIFFSHRDMNOVLDAYENKRPFLYLTG 114  
 DB 63 KLIRFGSSKIDKELIARFEKITGKPAHHFIRGMFESHDLHTLTLRQGKRPFLYLTG 122

DR MGD:MG1:104630; Wars.  
DR InterPro: IPR001412; tRNA-syn<sup>t</sup>-I.  
DR InterPro: IPR002306; tRNA-syn<sup>t</sup>-t<sup>trp</sup>.  
DR InterPro: IPR000738; WHEP-TRS.  
DR Pfam: PF00458; WHEP-TRS; 1.  
DR PRINTS: PR01039; tRNA-syn<sup>t</sup>-t<sup>trp</sup>.  
DR PROSITE: PS00178; AA-TRNA-LIGASE\_I; 1.  
SO SEQUENCE 475 AA; 5361 MW; C3467EE85521DE4C CRC64

Query Match	91.3%	Score 2051;	DB 11	Length 475;
Best Local Similarity	90.3%	Pred. No. 1.1e-171;		
Matches 383; Conservative	21;	Mismatches 20;		
				Indels 0; Gaps 0;

Qy	1	MSYRAAGEDYKADCPGPNPAPTSNHGDDATAEDEEDVDMPYQTSASAKIDYDKLVRF	60
Db	52	MSYRAAMEEYKACGCPGPNPTAGRNCDSDATKASEDVEDYDPTVYTSASAKIDYDKLVOR	111
Qy	61	GSSKIDKELINRIERATGQRPHEHLRRGIFFSHADMNQVLDAVENKRPFLYLTGRGSPSE	120
Db	112	GSSKIDKELINRIERATGQRPHEHLRRGIFFSHADMNQVLDAVENKRPFLYLTGRGSSSE	171
Qy	121	AMHGHILPFEFTWKLQDVFNWPLVIQMTDEKTLKMDLRLDQAYGGAIVENADITACGF	180
Db	172	AMHGHILPFEFTWKLQDVFNWPLVIQMSDEKTLKMDLRLQOAYSTVENAKDITACGF	231
Qy	181	DINKTFIISDDIYDGMSSGCFKYNVYKLOKHVTFNQVNGIFGFTSDICIGISPAIDQAP	240
Db	232	DINKTFIISDLEIYMGQSGFIRNYYVKLOKHVTFNQVNGIGFTSDICIGISPAVQAAP	291
Qy	241	SFSNSFPQIFPRDRTDIOCLIPCAIDODPYFPMTRDVAAPRIGYPRKALLHSTFPFALQAO	300
Db	292	SFSNSFPKIFPRDRTDIOCLIPCAIDODPYFPMTRDVAAPRIGHPRKALLHSTFPFALQAO	351
Qy	301	TKMSADPNSSSIFLTDPAKQIKTVKNHAFSGSGBDTIEHNQFSGNDVDVYSEPMYLFEL	360
Db	352	TKMSADPNSSSIFLTDPAKQIKTSVKNHAFSGSGBDTIEHNQFSGNEVDVSEMYLFEL	411
Qy	361	EDDDKLEQIRRDYTSAGALTEELKALALIEVLOPILAEHQARRKEVTDIEIYKEFMTPRKLS	420
Db	412	EDDDRLQIRRDYTSAGALTEELKKTLLIDVLOPLIAEQARRKAVTEIYKEFMTPRQLS	471
Qy	421	FDFO 424	
Db	472	FHFQ 475	

RESULT 2			
099J58			
ID 099J58	PRELIMINARY:	PRT:	481 AA.
AC 099J58;			
DT 01-JUN-2001 (TREMBLrel, 17, Created)			
DT 01-JUN-2001 (TREMBLrel, 17, Last sequence update)			
DT 01-DEC-2001 (TREMBLrel, 19, Last annotation update)			
DE TRIPTOPHANYL-TRNA SYNTHETASE.			
GN			
OS Mus musculus (Mouse).			
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus			
OX NCBI_TaxID=10090;			
[1]			
RP SEQUENCE FROM N.A.			
RP TISSUE=MAMMARY TUMOR. MAP-TAG MODEL. 5 MONTHS OLD, GROSS TISSUE.;			
RA Strausberg R.;			
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.			
DR EMBL: BC003450; AAH03450.1; -.			
DR MGI: MGI:104630; Wats.			
DR InterPro: IPR002305; tRNA-synt_1b.			
DR InterPro: IPR001412; tRNA-synt_1.			
DR InterPro: IPR002306; tRNA-synt_trp.			
DR InterPro: IPR000738; WHEP-TRS			
DR Pfam: PF00579; tRNA-synt_1b.1.			
DR Pfam: PF00458; WHEP-TRS; 1.			
DR PRINTS: PRO1039; TRNASYNTHPE			

DR PROSITE; PS00178; AA-TRNA-LIGASE-I; 1.  
DR PROSITE; PS00762; WHEP\_TRS; 1.  
KW Aminoacyl-tRNA synthetase.  
SQ SEQUENCE 481 AA; 54325 MW; A754E1DD5E82EF3 CRC64;

Query Match	91.1%	Score 2047	DB 11	Length 481
Best Local Similarity	90.1%	Pred. No. 2.5e-171		
Matches 382; Conservative	2;	Mismatches 20;	Indels 0;	Gaps 0

Qy	1	MSYKAAEEDYKACDPGNPAPTSNHGPDAAEAEDEVDPMPTVOTSASAKGIDDKLVRF	60
Db	52	VSYKAAEEDYKACGPRGNPTAGRNCDSDATKASEDVEDPMPTVTSASAKGIDDKLVRF	111
Qy	61	GSSKIDKELINRIERATGQREHNLRGIFPSHSDMNOVDVAYENKRPYLTYGRGSPSE	120
Db	112	GSSKIDKELINRIERATGQREHNLRGIFPSHSDMNOVDVAYENKRPYLTYGRGSPSE	171
Qy	121	AMHVHLLPFIETKKLODVENVPLVIGMTDEKTLMKDLTLDOAYGDAVENADITACGF	180
Db	172	AMHGLHVPFIETKKLODVENVPLVIGMTDEKTLMKDLTLDOAYGDAVENADITACGF	231
Qy	181	DINKTFIFSDLDYOMSSGFYKNVYIKQKHTFNOYVGIFGFTSDIGKISPPAIDAAP	240
Db	232	DINKTFIFSDLEYGMOQSGFRNVYIKQKHTFNOYVGIGFTSDIGKISPPAIDAAP	291
Qy	241	SFSNSFPQIFPHRDIDICLLPCALDIDODEYFMTDRDVAARIGYPRALLHSTFPFALOGAO	300
Db	292	SFSNSFPKIFPRDRDIDICLLPCALDIDODEYFMTDRDVAARIGHPRALLHSTFPFALOGAO	351
Qy	301	TKMSASDPNNSSTFLPDTAKOKTKTVNKNHAFSGSGDITIEHNQFEGNCDVUYVSFMYLFEFL	360
Db	352	TKMSASDPNNSSTFLPDTAKOKTKSVNKNHAFSGSGDITVEHNQFEGNCEVDVVSFMYLFEFL	411
Qy	361	EDDDLEOIRKDYTSGAMLTGELKALIEVLOPRLIAEHQARRKEVLTDEIYIEFMTPRKLS	420
Db	412	EDDDRLIEDIRKDYTSGAMLTGELKKTLLIDVLPRLIAEHQARRKAVTEBYUKKEFMTPRQLS	471
Qy	421	EDFQ 424	
Db	472	FHFQ 475	

RESULT 3			
Q9D8R9	PRELIMINARY;	PRT;	329 AA.
ID Q9D8R9			
AC Q9D8R9;			
DT 01-JUN-2001 (TREMblrel. 17, Created)			
DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)			
DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)			
DE TREPTOPHANYL-TRNA SYNTHETASE.			
GN WAKS.			
OS Mus musculus (Mouse).			
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX NCBI_TaxID=10090;			
[1]			
RP SEQUENCE FROM N.A.			
RC STRAIN=C57BL/6J; TISSUE=PANCREAS;			
RX MEDLINE=21085660; PubMed=11217851;			
RA Kawai J., Shinaawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,			
RA Akawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,			
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamamata I.,			
RA Salto T., Okazaki Y., Gobjori T., Bono H., Kasukawa T., Salto R.,			
RA Kadota K., Matsuda H.A., Ashurner M., Batalov S., Casavant I.,			
RA Fletschmann W., Gaasterland T., Gissi C., King B., Kochia H.,			
RA Kuell P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush			
RA Schimi L.M., Staubi F., Suzuki R., Tomita M., Wagner L., Washio			
RA Sakai K., Okido T., Furuno M., Aono H., Balderelli R., Barsh G.,			
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,			
RA Brownstein M.J., Hill D., Hofmann M., Hume D.A., Kamya M., Lee N.H.,			
RA Gustinich S., Hill D., Hofmann M., Hume D.A., Kamya M., Lee N.H.,			
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Kombaret P.,			
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,			

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 24, 2002, 12:49:45 ; Search time 22.9977 Seconds  
(without alignments)  
3189.442 Million cell updates/sec

Title: US-09-813-718-10\_COPY\_48\_471  
Perfect score: 2246  
Sequence: 1 MSYKAAAGEDYKADCPGPNP.....VTDEIVKEFWMPKRLSDFQ 424

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

SPREMBL\_19:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriap:\*  
17: sp\_archeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2051	91.3	475	11 Q9DC65	Q9DC65 mus musculus
2	2047	91.1	481	11 Q9J058	Q9J058 mus musculus
3	1629	72.5	329	11 Q9DBR9	Q9DBR9 mus musculus
4	1537	68.4	305	11 Q70184	Q70184 cavia porce
5	1382	61.5	430	5 Q9U4Y1	Q9U4Y1 drosophila
6	1379.5	61.4	420	5 Q9U4Y0	Q9U4Y0 drosophila
7	1378	61.4	430	5 Q9V8G2	Q9V8G2 drosophila
8	1301	57.9	402	10 Q9SR15	Q9SR15 arabidopsis
9	1027	45.7	324	5 Q9U1R2	Q9U1R2 caenorhabdi
10	910.5	40.4	406	17 Q976M1	Q976M1 sulfolobus
11	907	40.5	386	17 Q97ZX0	Q97ZX0 sulfolobus
12	800.5	35.6	490	5 Q9U1F5	Q9U1F5 leishmania
13	682	30.4	136	6 Q9S2S5	Q9S2S5 sus scrofa
14	626.5	27.9	301	17 Q9S9S4	Q9S9S4 pyrococcus
15	534.5	23.8	380	17 Q9H866	Q9H866 halobacteri
16	388.5	17.3	136	6 Q9TS88	Q9TS88 bos taurus

17	310	13.8	111	5 Q95YL8	Q95YL8 encephalito
18	299.5	13.3	426	17 Q978V8	Q978V8 thermoplasma
19	294	13.1	157	5 Q9U533	Q9U533 trypanosoma
20	286	12.7	109	4 Q9UD15	Q9UD15 homo sapien
21	274.5	12.2	513	17 Q9H883	Q9H883 halobacteri
22	268.5	12.0	426	17 Q9H1W5	Q9H1W5 thermoplasma
23	188	8.4	364	17 Q9YA64	Q9YA64 aeropyrum p
24	185.5	8.3	341	16 Q97N42	Q97N42 streptococc
25	182.5	8.1	340	16 Q9YX44	Q9YX44 streptococc
26	181.5	8.1	331	17 Q9Y9Z1	Q9Y9Z1 thermoplasma
27	174.5	7.8	341	16 Q9C0D1	Q9C0D1 lactococcus
28	167	7.4	351	16 Q9RVD6	Q9RVD6 delnecoccus
29	164.5	7.3	895	10 Q9SGN2	Q9SGN2 arabidopsis
30	163	7.3	460	10 P93018	P93018 arabidopsis
31	158.5	7.1	375	17 Q9V0Z7	Q9V0Z7 pyrococcus
32	154.5	6.9	102	1 Q07119	Q07119 halobacteri
33	152.5	6.8	682	5 Q9N9B8	Q9N9B8 leishmania
34	149.5	6.7	327	17 Q9H862	Q9H862 halobacteri
35	149.5	6.7	408	10 P93363	P93363 nicotiana t
36	146	6.5	375	17 Q58739	Q58739 pyrococcus
37	145	6.5	332	17 Q9HKT3	Q9HKT3 thermoplasma
38	142.5	6.3	528	11 Q91WQ3	Q91WQ3 mus musculu
39	139	6.2	294	17 Q96TV3	Q96TV3 sulfolobus
40	135	6.0	365	16 Q9KNV7	Q9KNV7 vibrio chol
41	134.5	6.0	525	5 Q9VW60	Q9VW60 drosophila
42	130.5	5.8	419	16 Q9ZBB1	Q9ZBB1 listeria in
43	129	5.7	339	2 Q9KZK7	Q9KZK7 streptomyce
44	124	5.5	344	16 Q9AC05	Q9AC05 caulobacter
45	123	5.5	347	10 Q82313	Q82313 arabidopsis

#### ALIGNMENTS

RESULT 1  
ID Q9DC65 PRELIMINARY: PRT: 475 AA.  
AC Q9DC65;  
DT 01-JUN-2001 (TREMBL:rel. 17, Created)  
DT 01-JUN-2001 (TREMBL:rel. 17, Last sequence update)  
DE 01-DEC-2001 (TREMBL:rel. 19, Last annotation update)  
BT ADULT MALE LUNG CDNA, RIKEN FULL-LENGTH ENRICHED LIBRARY,  
DE CLONE:1200002C07, FULL INSERT SEQUENCE.  
GN WARS.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N. A.  
RC SPRAIN=C57BL/6J; TISSUE=LUNG;  
RX MEDLINE=21085660; PubMed=11217851;  
RA Kawai T., Shinaigawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
RA Arikawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana I.,  
RA Saito T., Okazaki Y., Gotohori T., Bono H., Kasukawa T., Saito R.,  
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochia H.,  
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,  
RA Sacrali L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bona M.F.,  
RA Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M.,  
RA Gustincich S., Hill D., Hofmann M., Hune D.A., Kamiya M., Lee N.H.,  
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombauts P.,  
RA Nodone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
RA Saeki K., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,  
RA Wyszynski-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohzuki S.,  
RA Hayashizaki Y.,  
RT "Functional annotation of a full-length mouse cDNA collection.";  
RL Nature 409:685-690(2001).  
DR EMBL; AK004541; BAB23357.1; .



Query Match 7.7%; Score 172; DB 1; Length 394;  
Best Local Similarity 22.7%; Pred. No. 7.8e-07;  
Matches 88; Conservative 62; Mismatches 132; Indels 106; Gaps 22;

```
OY 94 RDMNOVL-----DAVE-NKKPFYLYTGRGSSSEAMHGHLPFTFKWLQDFNV--PL 144
   ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 18 KNIQEVLPNQIKDVLEVQKRHLKLYWGTPGRP-HCGYFVP--MTK-LADFLKAGEV 73
   ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
OY 145 VIQMTDDEKYL-----WKDLTLD---QATGDAVENAKDIIACGFDINKTF 186
   ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 74 TVLLADLHAFLDNMKAPLEVNVYRAKYYELLTKAILRSINVPDEKLFVVGSSYQLTDPY 133
   ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
OY 187 IFSDLDMGMSGGFYKNVYK-----IOKHVTENQYKIGFTDSDCIGKISPPAIQAAPS 241
   ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 134 ---TWDIFRLSNIVSQNDKRAKAGADVKKOYANPLLSGLI-----YPLMQA--- 175
   ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
OY 242 FSNSEFPQIFRDRPTDIOCLIPCAIDDPYFRMTRDVAAPRIGYKPPALHSTFPAL-QGAQ 300
   ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 176 -----LDEQFLVDYDCQG-GVDQKIFVLAENLPISLGYKKRAHLMNPVPGLAQGG- 226
   ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
OY 301 TKMSASDPNSSIFLDTAKQITKYKNAFSGSGRDTIEH-----RQFGG 345
   |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 227 -KMSASDPNSKIDLEEPQYKKKINSAPCSPG--NVEENGSLSFVOYVIAPIOELKFGT 283
   |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
OY 346 NCDVVSFWYLTFFLEDDK-----LEQIRKDYTSGAMLTGELK---KALIEVLQ 392
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 284 N-----HEEFIDRPEKFGGPITYKSFEEMKLAFKEEKLSPDCLKIGVADAINELLE 335
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
OY 393 PLIAEHQARKKEVTDEIVKEF--MTPRK 418
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 336 P-IRQEFANKEFQEAASEKGYVPATPOK 362
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
```

Search completed: October 24, 2002, 12:51:43  
Job time : 8.77387 secs

```

AC Q46127;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA ligase)
DE (TrpRS).
GN TRPS OR TRSA.
OS Clostridium longisporum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1523;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=B6405;
RA Brown G.D., Thomson J.A.;
RL Submitted (DEC-1995) to the EMBL/Genbank/DBJ databases.
CC -i- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +
CC diphosphate + L-tryptophanyl-tRNA(Trp).
CC -i- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -i- SUBCELLULAR LOCATION: Cytoplasmic.
CC -i- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; L49336; AAC05711.1; -.
DR HSSP; P00953; 1D2R.
DR InterPro; IPR002305; tRNA-synt_1b.
DR InterPro; IPR001412; tRNA-synt_1.
DR InterPro; IPR002306; tRNA-synt_tlp.
DR Pfam; PF00579; tRNA-synt_1b; 1.
DR PRINTS; PR01039; TRNASYNTTRP.
DR PROSITE; PS00178; AA-TRNA-LIGASE_I; 1.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding.
FT SITE 12 20 "HIGH" REGION.
FT SITE 201 205 "KMSKS" REGION.
FT BINDING 204 204 ATP (BY SIMILARITY).
SQ SEQUENCE 341 AA; 38256 MW; 692C820F5A08E4D1 CRC64;

Query Match 7.8%; Score 175.5; DB 1; Length 341;
Best Local Similarity 24.7%; Pred. No. 3; 3e-07;
Matches 78; Conservative 53; Mismatches 150; Indels 35; Gaps 11;

OY 111 LVTGRGSPSEAMVGHLPFIPTK-WLQDVFNVPLVIOQNTDDEKYLIMKDLTLDQAYGDVAV 169
DB 6 ILTGRPTGK-LHGHVYGLSKNRVQLOSGDYRSPIMADQALLDNARNPEKIRNSLI 64
OY 170 ENAKDILACGFINKFTIF--SDLDYMGSSGFYKNV---KIQKHVFNQYKGFPGFD 224
DB 65 EALDYLAAGIDPLKSTLIVQSQIPELNELTMHYLNLVLSLENNPYYKAEIKQKNFEN 124
OY 225 SDICIKISPPAIQAAPSFNSPFOIFRDRDIOCLPCALIDDPFPMRDVAPRIG--- 281
DB 125 STIPAGFLIYPSQADITR-----FKATP-----VPVGEDDLPMIEQARELYRSFNIT 173
OY 282 -----YKPRALHSTF--FPALQAGQTKMASADPNSSIFLTDTAQKQIKTKVKNHAFSG 333
DB 174 GKVELVERKAVIPKGTIGRLPCTDG-KAKMSKSTGN-AIYLADEADVIKQVMSMYTDPN 231
OY 334 RGTIEHNRFGGNCVDVSMFLTFLEDDDKLEQIRKRYTSGAMITGLKALLLEVDP 393
DB 232 KIKVYDPOGVEN---TVFTYLDLFCQKTEPLEMKAHYSGKGLGVKKVFLMEIIDA 287
OY 394 LIAEQARKEVTDEI 409
DB 288 ELEPIRNKRKEFOKDI 303

```

```

RESULT 15
ID SYC_YEAST STANDARD: PRT: 394 AA.
AC P36421.
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Tyrosyl-tRNA synthetase, cytoplasmic (EC 6.1.1.1) (Tyrosyl--tRNA
DE ligase) (TYRS)
GN TySI OR MGM104 OR YGR185C OR G7522.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=93286133; PubMed=8509419;
RA Chow C.M., Rajbhandary U.L.;
RT "Saccharomyces cerevisiae cytoplasmic tyrosyl-tRNA synthetase gene.
RT Isolation by complementation of a mutant Escherichia coli suppressor
RT tRNA defective in aminoacylation and sequence analysis.";
RL J Biol. Chem. 268:12855-12863(1993).
RN [2]
RP SEQUENCE FROM N.A.
RA Guan M.-X., Chen X.-J., Clark-Walker G.D.;
RL Submitted (MAY-1993) to the EMBL/Genbank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C;
RX MEDLINE=97279231; PubMed=9133739;
RA Arroyo J., Garcia-Gonzalez M., Garcia-Saez M.I., Sanchez-Perez M.,
RA Nombela C.;
RT "DNA sequence analysis of a 23,002 bp DNA fragment of the right arm
RT of Saccharomyces cerevisiae chromosome VII.";
RL Yeast 13:357-363(1997).
RN [4]
RP SEQUENCE OF 1-36 FROM N.A.
RC STRAIN=Bj926;
RX MEDLINE=95087887; PubMed=7995524;
RA Henry N.L., Campbell A.M., Fearer W.J., Poon D., Well P.A.,
RA Kornberg R.D.;
RT "TFII-F-TAF-RNA polymerase II connection.";
RL Genes Dev. 8:2868-2878(1994).
CC -i- CATALYTIC ACTIVITY: ATP + L-tyrosine + tRNA(Tyr) = AMP +
CC diphosphate + L-tyrosyl-tRNA(Tyr).
CC -i- SUBCELLULAR LOCATION: Cytoplasmic.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; L12221; AAB59329.1; -.
DR EMBL; X71998; -. NOT ANNOTATED_CDS.
DR EMBL; Z72870; CAA97211.1; -.
DR EMBL; X99074; CAA67529.1; -.
DR EMBL; U13015; AAA61641.1; -.
DR PIR; A45999; A45999.
DR SGD; S0003417; TYSI.
DR InterPro; IPR002305; tRNA-synt_1b.
DR InterPro; IPR001412; tRNA-synt_1.
DR InterPro; IPR002307; tRNA-synt_tyr.
DR Pfam; PF00579; tRNA-synt_1b; 1.
DR PRINTS; PR01040; TRNASYNTTHYR.
DR PROSITE; PS00178; AA-TRNA-LIGASE_I; FALSE NEG.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding.
FT SITE 48 56 "HIGH" REGION.
FT SITE 227 231 "KMSKS" REGION.
SQ SEQUENCE 394 AA; 44020 MW; 57EBDB9BE6D054B7 CRC64;

```

```

CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AP000664; BAA81476.1; -
DR InterPro; IPR002305; tRNA-synt_1b.
DR InterPro; IPR001412; tRNA-synt_1.
DR InterPro; IPR002306; tRNA-synt_trp.
DR Pfam; PF00579; tRNA-synt_1b; 1.
DR PRINTS; PR01039; TRNASYNTTRP.
DR PROSITE; PS00178; AA_TRNA_LIGASE_I; FALSE_NEG.
DR Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW Complete proteome.
FT SITE 81 89 "HIGH" REGION.
FT SITE 258 262 "KMSKS" REGION.
SQ SEQUENCE 374 AA; 42400 MW; A72635B7CA3F9189 CRC64;

Query Match 15.7%; Score 353; DB 1; Length 374;
Best Local Similarity 30.4%; Pred. No. 6,6e-22;
Matches 119; Conservative 70; Mismatches 165; Indels 38; Gaps 17;

QY 38 VDPWTVQVSSAKIDYDKLYRFGSSKIDKELINIERATGGRPHNPLRGFFSHRDMN 97
DB 8 LDPW---GAVEIKDYDLRLRTFGIRPSEVL--PLLRKAGMEPSFLMRKGIIFGHRD 61
QY 98 QVLDAENKRPFLYLTGRGSEAMHVGHLIPFTKMLQ-DVFNVPVLIOWTDEKILM 156
DB 62 KILEKAKGERAVALTGTMPGSK-FHFGHKLVLDLILOKNGFV--FVALADEAAV 118
QY 157 KDLTLDQAYGDAVEN-AKDIIACGFINKT-FIFSDLYMGMSGFYKVVYKIQHVTEN 214
DB 119 RRIREEAVRIAYEYIYANMIALGLDPKDTREYFQ---TNGPTPYRLIQLFSGKVTA 174
QY 215 QVKGIFG-FTPSDCIGKISFPAIOAAPSNSFPQIFRDRIQCLICALDDDDPYFMT 273
DB 175 EMEALYIGELTPAKKMASLT---QAADILHVQLDYGGYR--HVVVPGADQDPHRLT 227
QY 274 RDVAPR---IGYRPALLHSTFFPALOGAOTKMSASDPNSIFLTPRAKQIKTVKNKA 329
DB 228 RDLADRMAGVVELERPATYHKLPGLDQ--RKMSSTRDSTIFLTPDEPAKMKLFR-A 284
QY 330 FSGGDTIEHRQFGGNCV-DVSFMYLTFLLEDKLEQIRKDYTS---GAMLTGELIK 385
DB 285 LTGGRATAEQRRILGVEPVCSVYHMDYLHMPDGEVKNH---YTSCLKILGCECKQ 341
QY 386 ALIEVLOPLIAHQARRKEVNDIYKEFMTPR 417
DB 342 IAWEKLERFLAEHOSLERKAKTIAMKLVPPER 373

RESULT 13
SYN_ARCFU STANDARD; PRT; 323 AA.
AC 029482;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Tyrosyl-tRNA synthetase (Ec 6.1.1.1) (Tyrosine--tRNA ligase) (TYRS).
GN TYRS OR AF0776.
OS Archaeoglobus fulgidus.
OC Archaea; Euryarchaeota; Archaeoglobales; Archaeoglobaceae;
OC Archaeoglobus.
OX NCBI_TaxID=2234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=98049343; PubMed=9389475;

```

```

RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA Kechum K.A., Dodson R.J., Gwin M., Hickey E.K., Peterson J.D.,
RA Richardson D.L., Kerlavage A.R., Graham D.E., Kyriades N.C.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
RA Cotton M.D., Spillgas T., Artlich P., Kaine B.P., Sykes S.M.,
RA Sadow P.W., d'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RA "The complete genome sequence of the hyperthermophilic, sulphate-
RT reducing archaeon Archaeoglobus fulgidus."
RL Nature 390:364-370(1997).
CC -1- CATALYTIC ACTIVITY: ATP + L-tyrosine + tRNA(Tyr) = AMP +
CC diphosphate + L-tyrosyl-tRNA(Tyr).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE001051; AAB90462.1; -
DR TIGR; AF0776; -
DR InterPro; IPR002305; tRNA-synt_1b.
DR InterPro; IPR001412; tRNA-synt_1.
DR InterPro; IPR002307; tRNA-synt_lyr.
DR Pfam; PF00579; tRNA-synt_1b; 1.
DR PRINTS; PR01040; TRNASYNTTRP.
DR PROSITE; PS00178; AA_TRNA_LIGASE_I; 1.
DR Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW Complete proteome.
FT SITE 41 49 "HIGH" REGION.
FT SITE 214 218 "KMSKS" REGION.
SQ SEQUENCE 323 AA; 36616 MW; A655ABE4A5116642 CRC64;

Query Match 8.5%; Score 192; DB 1; Length 323;
Best Local Similarity 23.1%; Pred. No. 1.3e-08;
Matches 78; Conservative 67; Mismatches 127; Indels 66; Gaps 15;

QY 92 SHRMNVLDAYENKRPFLYLTGRGSEAMHVGHLIPFTKMLQDVFNVPVLIOWTDE 151
DB 19 TEEELRLQLETKRPR--AYGYEPGE-IHLGHMTVQKLMDEA-GFEIIVLADI 73
QY 152 EKYLMKDLTLDQAYGDAVENAKDIACGFDINKFTIFSDLYMGMSGFYKVVYKIQHV 211
DB 74 HAYLNEKTFPEELIAEVADYKNNKVFALGLDSRAKFLVGSYQ--LSRYIVLDYKMAIT 132
QY 212 TFNOYK---GIFGTFSDCIGKISFPAIOAAPSNSFPQIFRDRIQCL-IPCA--- 263
DB 133 TLNRARRSMDEVSRKKEPDMVSQMIYPLMQA-----LDIALHGLDYLAVGG 177
QY 264 IDQDPYFMRTDVAPRTGYPKPALLHSTFFPALOGAOTKMSASDPNSIFLTPRAKQIK 323
DB 178 IDORKIHMLARENPLRGYSPPVCLHPILVGLDQ--OKMSSSKN-YISVDPPEEVER 234
QY 324 KVNK-HAFSG-----GRDTIEHRQFGGNCV-DVSFMYLTFLLEDKLEQIRKDYTS 363
DB 235 KIRKAYCPAGVVEENPILDIKAYHLLPFGKIYVERDAKFGG--DVE---YASF----- 283
QY 364 DKLEQIRKDYTSGAMLTGELKALIEVLOPLIAHQAR 401
DB 284 ---EELADPFGSGQLHPDLKIAVAKYLMMLLEDARRK 318

RESULT 14
SYN_CIOLO STANDARD; PRT; 341 AA.
ID SYN_CIOLO

```



```

FT NON_TREE 134 134
SQ SEQUENCE 134 AA; 15744 MW: 60E2935B7E1E344F CXC64;

Query Match 17.2%; Score 386; DB 1; Length 134;
Best Local Similarity 51.5%; Pred. No. 2.9e-25;
Matches 68; Conservative 25; Mismatches 35; Indels 4; Gaps 1;

QY 35 EDFVPMWTVQTSASK-----GIDYKDLIYFSGSKIDKELINIERATGQRPHPFRRGIF 90
   1 : | | : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 3 EQRITPMDVEVSTDEVPAYDIDYKDIINQFGCEKRNQALADLEKLSCKPAHYFFRRGIV 62
   1 : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 91 FSHRDMNOVLDAENKKPFYLYTGRGSPSEAMVGHLPFIETFKVLQDVFNPLYIQMTD 150
   1 : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 63 FAHRPFNLLDEIANNRPFFIYTGSGPSKMTWIGHTITPFLCKRYMDAFKIRLVIQITD 122
   1 : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 151 DEKYLMDLTLD 162
   1 : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 123 DEKFLMKSMRL 134

RESULT 11
SYW_ARCFU
ID SYW_ARCFU STANDARD: PRT: 420 AA.
AC 028579;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA ligase)
DE (TrpRS).
GN TRPS OR AF1694.
OS Archaeoglobus fulgidus.
OC Archaeae; Euryarchaeota; Archaeoglobales; Archaeoglobaceae;
OC Archaeoglobus.
OX NCBI_TaxID=2234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=98049343; PubMed=9389475;
RA Kleink H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA Kechum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
RA Richardson D.L., Kervilave A.R., Graham D.E., Kyrides N.C.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Kinkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodet A., Zhou L.,
RA Overbeek R., Goeyne J.D., Weidman J.F., McDonald L., Uteckback T.,
RA Colton M.D., Spriggs T., Artlich P., Kaine B.P., Sykes S.M.,
RA Sadov P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RT "The complete genome sequence of the hyperthermophilic, sulphate-
RT reducing archaeon Archaeoglobus fulgidus."
RL Nature 390:364-370(1997).
CC -1- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +
CC diphosphate + L-tryptophanyl-tRNA(Trp).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/).
CC or send an email to license@sib-sib.ch).
CC
CC -----
CC EMBL; AE000986; AAB89554.1; -
CC TIGR; AF1694; -
CC InterPro: IPR002305; tRNA-synt_1b.
CC InterPro: IPR001412; tRNA-synt_1.
CC InterPro: IPR002306; tRNA-synt_1tp.
CC Pfam; PF00579; tRNA-synt_1b; 1.
CC PRINTS; PRO1039; TRNASYNTTRP.
CC PROSITE; PS00178; AA_TRNA_LIGASE_I; FALSE NEG.

```

Query Match	16.5%	Score 370.5	DB 1	Length 420
Best Local Similarity	27.3%	Pred. No. 2.7e-23		
Matches 118	Conservative 60	Mismatches 166	Indels 89	Gaps 12
Query	38	VDPVAVQTSASGAGIDYDKLIVFGSGSIDKELINIRERATGQRPHHFLRGFFSRRDMN	97	
Db	3	VTFPVEEGV----IDYSKLEEFQMGQPF-SEVLPELD-----NPHILMRGAGLFGHRDYM	52	
Qy	98	QVLDAVENKKRPVLYLTGKGPSSSEAMHVGHLIPFTFKWLQDVFNVPVLYOMTDDKYLWK	157	
Db	53	RILEAMOKKEPAAVNSGFMPSG-LPHFGHKMTDEIVHQSAGKAFV-AIDMEAHSVR	110	
Qy	156	DITLDOAGDAVENAKDITACGDFINKFTFESDDLYMGSSGFYKNV-KIQKHVFNQV	216	
Db	111	GLSWKTRRELGLMYKISIIALGRDAVLYFQS-----KSSHVXDLAFELSAEYNFSEL	164	
Qy	217	KCIFETSDSDICGISFPDIAQAPSPNSFPQIFRDRDITQCLICADIDQVFFMYTRPV	276	
Db	165	KATYGFNSDTSLAKMFVYATQADIL--HPQLSDFGCKPVVVEGADQDPHMRITRDL	221	
Qy	277	APRI-----	280	
Db	222	AARISFSEPEYEGGVRRSRKGAEYLSLRLEFPDKIYEHMIDIFGAEIERAVRKI	281	
Qy	281	-----GYPRALLHSTFFPALQGAQTKMSADPNSSIFETDTAKQIKTKYNKHAFFSG	333	
Db	282	EVEIGGAFAPFPSSYTHRETTGLTGG--KMSSSKPESSYISLDPPEEGAKKYMK-AFTGG	338	
Qy	334	RPTIEHROFGNCQVDVSEFMVLTFFLED-DOKLQIRKDYNSGMALGELKALILEYQ	392	
Db	339	RATAEORRLGGEPRDCVVEFLYSFHLIDSDDELQOIEACREGRGLCGCKKMAAEIVK	398	
Qy	393	PLIAEHOARKREV	405	
Db	399	SFLKEHQEKMEAV	411	
RESULT 12				
SYN_AERPE	ID	SYN_AERPE	STANDARD:	PRT: 374 AA.
AC	09Y924:			
DT	30-MAY-2000	(Rel. 39, Created)		
DT	30-MAY-2000	(Rel. 39, Last sequence update)		
DT	16-OCT-2001	(Rel. 40, Last annotation update)		
DE	TRYPLOPHANYL-tRNA synthetase (EC 6.1.1.2) (Tryptophan-tRNA ligase) (TTPRS).			
DE	GN	TTPS OR APE2461.		
OS	Aeropyrum pernix.			
OC	Archaea; Crenarchaeota; Desulfurococcales; Desulfurococcaceae;			
OC	Aeropyrum.			
OX	NCBI_TaxID=56636;			
OX	NCBI_TaxID=56636;			
OX	NCBI_TaxID=56636;			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-K1:			
RX	MEDLINE=99310339; PubMed=10382966;			
RA	Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,			
RA	Jin-no K., Takahashi M., Sekine M., Baba S.-I., Aikai A., Kosugi H.,			
RA	Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,			
RA	Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,			
RA	Yamazaki J., Kushiida N., Oguchi A., Aoki K.-I., Kudoh K.,			
RA	Nakamura Y., Nomura N., Sako Y., Kikuchi H.;			
RT	"Complete genome sequence of an aerobic hyper-thermophilic			
RT	crenarchaeon, Aeropyrum pernix K1."			
RL	DNA Res. 6:83-101(1999).			
CC	-i- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(TTP) = AMP +			
CC	diphosphate + L-tryptophanyl-tRNA(TTP).			
CC	-i- SUBCELLULAR LOCATION: Cytoplasmic.			

Db 113 RNMSPTTKELALNEYINIALGLDPEKINVIYLSQKYQV-----KDLALILSKRTMMS 167  
QY 215 QVKGIFGTTSDSCIGKISFPALQAPSFNSFPQIFRDRTP--DIOCLIPCAIDDDPYTRM 272  
Db 168 EMKAIYFGKFGKGTINIVAPYIVADIL--HPOLEDMISPEPKVVPVPGIDODPHRL 224  
QY 273 TRDVAPR---IGYPRPALHSTFPALQAGTQKMSASDPNSIFLTPAKOIKTKTVKNKA 329  
Db 225 TRDIANRKKERKFLPPSSITHRFMTGLGG--KMSSSKPEFAIFLTDDEKTVKKKIFS-A 281  
QY 330 FSGGRDITIEHRQFEG---NCDVDVSFMYLTFEEDDKLQIRKDYTSGAMLTGELKKA 386  
Db 282 KTGRETLEEHKKYGVPECEVVELFLY--HLIIDKELAEIYQKCSGELTGCKCKM 339  
QY 387 LIEVLIQPLIAEHQARKKEVTEIYK 411  
Db 340 AYERVEFLKDKERQAKELAVK 364

## RESULT 9

SYM\_METH  
ID SYM\_METH STANDARD: PRT: 364 AA.  
AC 026352;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA ligase)  
DE (TPRS).  
GN TRPS OR MTH251.  
OS Methanobacterium thermoautotrophicum.  
OC Archaea; Euryarchaeota; Methanobacteriales; Methanobacteriaceae;  
OC Methanothermobacter.  
OX NCBI\_TaxID=145262;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-DELTA H;  
RX MEDLINE=98037514; PubMed=9371463;  
RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,  
RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,  
RA Harrison D., Hoang L., Keagle P., Lumm W., Potlter B., Qiu D.,  
RA Spadatore R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,  
RA Jivani N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,  
RA McQuail S., Shimer G., Goyal A., Pietrowski S., Church G.M.,  
RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;  
RT "Complete genome sequence of Methanobacterium thermoautotrophicum  
RT strain: functional analysis and comparative genomics.";  
RL J. Bacteriol. 179:7135-7155(1997).  
CC -1- CATALYTIC ACTIVITY: ATP + L-Tryptophan + tRNA(Trp) = AMP +  
CC diphosphate + L-tryptophanyl-tRNA(Trp).  
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: AE000812; AAB84757.1; -  
DR InterPro: IPR002305; tRNA-synt\_Lb.  
DR InterPro: IPR001412; tRNA-synt\_L.  
DR InterPro: IPR002306; tRNA-synt\_Lb.  
DR Pfam: PF00579; tRNA-synt\_Lb; 1.  
DR PRINTS: PRO1039; TRNASYNTTRP.  
DR PROSITE: PS00178; AA\_TRNA\_LIGASE\_I; FALSE NEG.  
KW Aminoacyl-tRNA synthetase; protein biosynthesis; ligase; ATP-binding;  
KW Complete proteome.  
FT SITE 70 78 "HIGH" REGION.  
FT SITE 251 255 "KMSKS" REGION.  
SQ SEQUENCE 364 AA; 41301 MW; C2F348903338F61D CRC64;

Query Match 17.7%; Score 397.5; DB 1; Length 364;  
Best Local Similarity 27.9%; Pred. No. 1.3e-25;  
Matches 107; Conservative 76; Mismatches 153; Indels 47; Gaps 14;  
QY 38 VDPMTVOTSSAKGIDYDKLIVFGSSKIDKELINIRERATGCRPHFLRGIFFSHRDM 97  
Db 2 IDPW-----GSAK-LEYQDLINFGVRP-SEVLEDEV-----PEFSWLMRRGILIGHROYE 50  
QY 98 QVDAIYENKKRFFPYLYTGGSPSEAMHGHILPFTKMLQOVFNVPVLYQMTDDEKYLK 157  
Db 51 RIISAMKGEDEFVAVTGMPGSR-MHIGKMIIVOLRW-YRMAEAFIFPADMEAVSAR 108  
QY 158 DLTLDQAGDVLEN-ANKDIACGPDINK-----TFJSPDDYWGMSGFFKYNVK 206  
Db 109 GVDFEDSRRIAIIEIYIAGYIALGLDLEKDNHIVLYQSNLMVEDLAY----- 156  
QY 207 IQKHVFNQVKGIFGTTSDSCIGKISFPALQAPSFNSFPQIFRDRTPDIOCLIPCAIDQ 266  
Db 157 LAGKVNFEIARLYGFTGSTMAHMYAPILIOVDILHPQLDELGPR---PVIYVPGPDQ 213  
QY 267 DPYRMTDVAPRI---GYPRPALHSTFPALQAGTQKMSASDPNSIFLTPAKOIK 322  
Db 214 DPHRLTRDIAARFRDRYGFILPSTYHFMGGLTGG--KMSNRKPSAIFLSDTPEAE 271  
QY 323 TKVKNHAFSGRDTIEEHRQFEGNCDVDVSFMYLTFEEDDKLQIRKDYTSGAMLTG 381  
Db 272 AKI-RNATGREGELKEKRELGVPEECITETLLIHMSGSDSRLEETECRNGTLMCG 330  
QY 382 ELKRALIEVLQPLIAEHQARKKE 404  
Db 331 ECKNNTAEFIKKFPEELSVKREK 353

## RESULT 10

SYM\_ENCCU  
ID SYM\_ENCCU STANDARD: PRT: 134 AA.  
AC 096771;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA ligase)  
DE (TPRS) (Fragment).  
OS Encephalitozoon cuniculi.  
OC Eukaryota; Microsporidia; Unikaryoniidae; Encephalitozoon.  
OX NCBI\_TaxID=6035;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC MEDLINE=98277683; PubMed=9615449;  
RA Peyretailade E., Broussolle V., Peyret P., Metenier G., Gouy M.,  
RA Vlaeares C.P.;  
RT "Microsporidia, amitochondrial protists, possess a 70-kDa heat shock  
RT protein gene of mitochondrial evolutionary origin.";  
RL Mol. Biol. Evol. 15:683-689(1998).  
CC -1- CATALYTIC ACTIVITY: ATP + L-Tryptophan + tRNA(Trp) = AMP +  
CC diphosphate + L-tryptophanyl-tRNA(Trp).  
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).  
CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: AJ012470; CAA10034.1; -  
DR HSP: P00952; INYC.  
DR InterPro: IPR001412; tRNA-synt\_L.  
DR PROSITE: PS00178; AA\_TRNA\_LIGASE\_I; 1.  
KW Aminoacyl-tRNA synthetase; protein biosynthesis; ligase; ATP-binding.  
FT SITE 89 98 "HIGH" REGION.

```

SYW_PYRAB
ID SYW_PYRAB STANDARD; PRT; 385 AA.
AC 09UT11;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA ligase)
DE (TrpRS).
GN TRPS OR PAB1111.
OS Pyrococcus abyssi.
OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.
OX NCBI_TaxID=29292;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ORSAY;
RA Hellig R.;
RT "Pyrococcus abyssi genome sequence: insights into archaeal chromosome
   structure and evolution.";
RL Submitted (JUL-1999) to the EMBL/Genbank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +
   diphosphate + L-tryptophanyl-tRNA(Trp).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AJ248288; CAB50601.1; -
DR InterPro: IPR002305; tRNA-synt_1b.
DR InterPro: IPR001412; tRNA-synt_1.
DR InterPro: IPR002306; tRNA-synt_trp.
DR Pfam: PF00579; tRNA-synt_1b; 1.
DR PRINTS: PR01039; TRNASYNTTRP.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW Complete proteome.
FT SITE 82 "HIGH" REGION.
FT SITE 253 "KMSKS" REGION.
SQ SEQUENCE 385 AA; 45100 MW; 4C29D01414976B12 CRC64;

Query Match 35.8%; Score 803; DB 1; Length 385;
Best Local Similarity 45.3%; Pred. No. 2.8e-59;
Matches 173; Conservative 63; Mismatches 120; Indels 26; Gaps 9;

QY 35 EDF-VDPWTVOTSSAKGIDYDKLIVRFSSSKIDKELINRIERATGQRPHHFLRGIFFSH 93
   ||| ||| : ||| ||| : ||| ||| : ||| ||| : ||| |||
DB 3 EDFKVTPEVEGV---VDYNKLIEHFGTSPLEELKELTALTSSELPLEFRKRFSSH 58
   ||| ||| : ||| ||| : ||| ||| : ||| ||| : ||| |||
QY 94 RDMNOVDAYENKKPFYLYTGRGSSSEAMHGHILPFTFKWLQDVNPVLYIQMTDEK 153
   ||| ||| : ||| ||| : ||| ||| : ||| ||| : ||| |||
DB 59 RDYDVLQDYEGRGFFLYTGRGPGS-PMHGHILIPFATKWLQKFGVNIYIQTTDEK 117
   ||| ||| : ||| ||| : ||| ||| : ||| ||| : ||| |||
QY 154 YLMKD-LTLDAQYDAVNANAKDIACGPDINKTFESDLDMGSSGKYKNVYKQIKVY 212
   ||| ||| : ||| ||| : ||| ||| : ||| ||| : ||| |||
DB 118 FLEKELNLFEDTKHAYENLIDIIAVGPDPTFTFQNSE---TKLYEMALPIAKKIN 173
   ||| ||| : ||| ||| : ||| ||| : ||| ||| : ||| |||
QY 213 FNOVVGIGFTLSDICIGKISFPAIDAAFSFNSFQIFRDTIDICLIPCAIDQDPYFRM 272
   ||| ||| : ||| ||| : ||| ||| : ||| ||| : ||| |||
DB 174 FSMANAVGFTFQSKIGMIFPAIDIAPTF-----FEKR---KCLIPALADQDPYKRL 223
   ||| ||| : ||| ||| : ||| ||| : ||| ||| : ||| |||
QY 273 TRDVAIRIGYKPAIIHSTFEPALOGAOTKMSASDPNSSIFLTDTAKOIKTKVKNHAFS 332
   ||| ||| : ||| ||| : ||| ||| : ||| ||| : ||| |||
DB 224 QRDFAESGYYTAIHSKFVSLTSLGKMSASKEPATAIYLTDSPEDEVKKVMWFALTG 283
   ||| ||| : ||| ||| : ||| ||| : ||| ||| : ||| |||
QY 333 GRDITEEHROGNCNDVQVSEFWYLFLEDDDKLEQIRKY---TSGMLTGEIKKALIE 389
   ||| ||| : ||| ||| : ||| ||| : ||| ||| : ||| |||
DB 284 GRPTLEOREKREKGEPEKVFEMLEIFFEEDK--KLKERYACKNGELTGCECKRYLIS 341
   ||| ||| : ||| ||| : ||| ||| : ||| ||| : ||| |||
QY 390 VLOPLIAHQARKEVTEIYK 411

```

```

DB 342 KIOFLKHEKQKRRKKAERQIEK 363
   : | : ||| ||| : | |
RESULT 8
SYW_METJA
ID SYW_METJA STANDARD; PRT; 370 AA.
AC 058810;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA ligase)
DE (TrpRS).
GN TRPS OR MJ1415.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;
OC Methanococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=6688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reisch C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhman J.L., Nguyen D.,
RA Uutterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Colton M.D., Roberts C.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klink H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
   jannaschii.";
RL Science 273:1058-1073(1996).
CC -1- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +
   diphosphate + L-tryptophanyl-tRNA(Trp).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U67882; AAB99425.1; -
DR TIGR: MJ1415; -
DR InterPro: IPR002305; tRNA-synt_1b.
DR InterPro: IPR001412; tRNA-synt_1.
DR InterPro: IPR002306; tRNA-synt_trp.
DR Pfam: PF00579; tRNA-synt_1b; 1.
DR PRINTS: PR01039; TRNASYNTTRP.
DR PROSITE: PS00178; AA-TRNA_LIGASE_I; FALSE NEG.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW Complete proteome.
FT SITE 75 "HIGH" REGION.
FT SITE 255 "KMSKS" REGION.
SQ SEQUENCE 370 AA; 42660 MW; E6C71107CF82B59D CRC64;

Query Match 18.2%; Score 409.5; DB 1; Length 370;
Best Local Similarity 30.9%; Pred. No. 1.3e-26;
Matches 119; Conservative 72; Mismatches 153; Indels 41; Gaps 16;

QY 40 PWTVOTSSAKGIDYDKLIVRFSSSKIDKELINRIERATGQRPHHFLRGIFFSHRDMNV 99
   ||| ||| : ||| ||| : ||| ||| : ||| ||| : ||| |||
DB 8 PW-ETPAV--IDYKKTMEQFGVKPIVDVLDLKEE-----HHFFRRNIIILGHRDFERI 57
   ||| ||| : ||| ||| : ||| ||| : ||| ||| : ||| |||
QY 100 LDAYENKKPFYLYTGRGSSSEAMHGHILPFTFKWLQ---DVFNVPVLYIQMTDEKILM 156
   ||| ||| : ||| ||| : ||| ||| : ||| ||| : ||| |||
DB 58 VDAIKNNKEFAVSGMPSGK-MHGHRMNVVDLKYOKYTDNINIPt---ADLEAWA 112
   ||| ||| : ||| ||| : ||| ||| : ||| ||| : ||| |||
QY 157 KDLTLDAQYDAV-ENAKDIIACGPDINKTFESDLDMGSSGKYKNVYK-OKHYTFN 214

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: Z50142; CA90500.1; -  
DR InterPro: IPR002305; tRNA-synt\_1b.  
DR InterPro: IPR001412; tRNA-synt\_1.  
DR InterPro: IPR002306; tRNA-synt\_1b.  
DR Pfam: PF00579; tRNA-synt\_1b; 1.  
DR PRINTS: PR01039; tRNA-synt\_1b.  
DR PROSITE: PS00178; AA-TRNA-LIGASE.1; 1.  
KW Hypothetical protein; Aminoacyl-tRNA synthetase; Protein biosynthesis;  
KW Ligase; ATP-binding.  
FT SITE 91 100 "HIGH" REGION.  
FT SITE 275 279 "KMSK" REGION.  
SQ SEQUENCE 395 AA; 44910 MW; E656AEB876C5FD9 CRC64;  
Query Match 53.9%; Score 1210; DB 1; Length 395;  
Best Local Similarity 59.6%; Pred. No. 4.3e-93;  
Matches 235; Conservative 60; Mismatches 91; Indels 8; Gaps 4;  
QY 34 EDEVDPTVOTV-----SAKGIDYDKLIVRGSSKIDKELINRIERATGQRPHFLRGI 89  
DB 4 EEQIYTPMDVAGSIYDGEKIDYERLIVQGTGRTKTPGEQIEREKLGKPKHLLRRGA 63  
QY 90 FFHSRDMNOVDAYENKRPFLYLTGGRPSSEAMHGHILPFIPTKMLQDVFNPLVDMT 149  
DB 64 FFSHDPEFMIDRYEQKKRPFYLTGGRPSDSMHLGHMIPFECMKMLQDVFNPLVDMT 123  
QY 150 DDEKFLMKD-LTLDOAYGDAVENAKDIACGFDINKTFIFSLDPMGSSGFFKVVVLIQ 208  
DB 124 DDEKFLKQGSVLEDCQGFARENADITAVGFDPKTITFNNSTIYG--GAFYQVNVKIA 181  
QY 209 KHTVENYKGIFFGTDSDICIGISFPALQAAAPFSNSFPQIFRDRTDIOCLIPCAIDDP 268  
DB 182 KCIANQSKACGFFDSDISICKIHFAISIQAAAPFSNSFPHFNGAKDIPCLIPCAIDDP 241  
QY 269 YFRMTDVAAPRIGYPRKPLHSTFPALQGAQTKMSASDPNSITFLTAAQIKTKVKNH 328  
DB 242 YFRMTDVAAPRIGYPRKPLHSTFPALQGAQTKMSASDPNSITFLTAAQIKTKVKNH 301  
QY 329 AFSGGRTIEEHRQFGNCQDVSPMYLTFFLEDDEKLEQIRKDYSGAMLGELKAL 388  
DB 302 AFSGGRTIEEHRQFGNCQDVSPMYLTFFLEDDEKLEQIRKDYSGAMLGELKAL 361  
QY 389 EVLQPLAEHQARKEVTDEIVKEFT-PRKLSF 421  
DB 362 KLLQOFVSDFOARSKVDEATLDMFMDSRKLEW 395  
RESULT 6  
SYNC\_YEAST STANDARD; PRT; 432 AA.  
ID SYNC\_YEAST STANDARD; PRT; 432 AA.  
AC 012109;  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DE tryptophanyl-tRNA synthetase, cytoplasmic (EC 6.1.1.2) (Tryptophan--  
DE tRNA ligase) (Tfprs).  
GN WRS1 OR YOL097C OR HRA32.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
OX NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96076631; PubMed=7502582;  
RA Vandenbol M., Durand P., Portetelle D., Hilger F.;  
RT "Sequence analysis of a 44 kb DNA fragment of yeast chromosome XV

RT including the Ty1-H3 retrotransposon, the sufl(+) frameshift  
RT suppressor gene for tRNA-Gly, the yeast transfer RNA-Thr-1a and a  
RT delta element.";  
RN Yeast 11:1069-1075(1995).  
RN [2]  
RP FUNCTION.  
RX MEDLINE=97197969; PubMed=9046085;  
RA John T.R., Ghosh M., Johnson J.D.;  
RT "Identification and expression of the Saccharomyces cerevisiae  
RT cytoplasmic tryptophanyl-tRNA synthetase gene.";  
RL Yeast 13:37-41(1997).  
CC -1- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Tyr) = AMP +  
CC diphosphate + L-tryptophanyl-tRNA(Tyr).  
CC -1- SUBUNIT: HOMODIMER.  
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: Z48149; CA88164.1; -  
DR EMBL: Z74839; CA99110.1; -  
DR SGD: S0005457; WRS1.  
DR InterPro: IPR002305; tRNA-synt\_1b.  
DR InterPro: IPR001412; tRNA-synt\_1.  
DR InterPro: IPR002306; tRNA-synt\_1.  
DR Pfam: PF00579; tRNA-synt\_1b; 1.  
DR PRINTS: PR01039; tRNA-synt\_1b.  
DR PROSITE: PS00178; AA-TRNA-LIGASE.1; 1.  
KW Aminoacyl-tRNA synthetase; protein biosynthesis; Ligase; ATP-binding.  
FT SITE 111 120 "HIGH" REGION.  
FT SITE 295 299 "KMSK" REGION.  
SQ SEQUENCE 432 AA; 49350 MW; C408F16973769736 CRC64;  
Query Match 51.8%; Score 1163; DB 1; Length 432;  
Best Local Similarity 54.8%; Pred. No. 4e-89;  
Matches 217; Conservative 71; Mismatches 100; Indels 8; Gaps 3;  
QY 30 ATEAEDEVDPMTV-----QTSASAGIDYDKLIVRGSSKIDKELINRIERATGQRPHF 84  
DB 19 STDVKEQVVTVDVGVGDEQGRAGNIDYDKLQFGTRPVNEETLKRFGQVGTGREPHF 78  
QY 85 LRGGFFSHRDMNOVDAYENKRPFLYLTGGRPSSEAMHGHILPFIPTKMLQDVFNPL 144  
DB 79 LRGGFFSERDFTKLLDYEQKRPFLYLTGGRPSDSMHLGHMIPFVETKMLQDEVFVPL 138  
QY 145 VIOMTDDKRYLWK-DLTLDOAYGDAVENAKDIACGFDINKTFIFSDIDYMGSSGFFKN 203  
DB 139 VIELTDDKFLFKHKLITINDYKNFARENADITAVGFDPKTITFNNSTIYG--GAFYET 196  
QY 204 VKIQRHVTENOVKGIFFGTDSDICIGISFPALQAAAPFSNSFPQIFRDRTDIOCLIPCA 263  
DB 197 VVRVSRQITGSTAKAVFGFNDSDICIGFHFASISQIAAFPSFNVALLDPKTPCLIPCA 256  
QY 264 IDQDYEFRMTDVAAPRIGYPRKPLHSTFPALQGAQTKMSASDPNSITFLTAAQIKTK 323  
DB 257 IDQDYEFRMTDVAAPRIGYPRKPLHSTFPALQGAQTKMSASDPNSITFLTAAQIKTK 316  
QY 324 KVNKAFFSGRDTIEEHRQFGNCQDVSPMYLTFFLEDDEKLEQIRKDYSGAMLGEL 383  
DB 317 KINKAFFSGGVSAVLHLELGGNPDVAVAYQLSFCKDDVFLKCYKYSGELLSEGM 376  
QY 384 KKALIEVQLPLAEHQARKEVTDEIVKEFTPRKL 419  
DB 377 KKCIETIQEFYKAFQERRAQVDEFTLKFVPHRL 412  
RESULT 7

```

|||||
Db 173 MHVGHILPFIETKWLQADVFDVPLVYMOSDDEKYLKKDLTLEQVGYGTLENKADIACGFD 232
Oy 182 INKTFIFSDLDYMGMSGSGFKYKVVVKIOKHVFNQVKGIFGFDSDCICKISPPAIQAAP 241
Db 223 VNKTFIFSDLDYMGMSGSGFKYKVVVKIOKHVFNQVKGIFGFDSDCICKISPPAIQAAP 292
Oy 242 FSNSPFOIFRDRDTIOCLIPCAIDODPYFRMRDVAAPRIGYKPPALHSTFFPALQAQ 301
Db 293 FSNSPFOIFRDRDTIOCLIPCAIDODPYFRMRDVAAPRIGYKPPALHSTFFPALQAQ 352
Oy 302 KMSADPNSSIFLDTAKQIKTKVKNKHAFFSGGRDTIEEHROFGGNCDDVDSFMYLTFE 361
Db 353 KMSADPNSSIFLDTAKQIKTKVKNKHAFFSGGRDTIEEHROFGGNCDDVDSFMYLTFE 412
Oy 362 DDDKLEQIRKDYSGAMLTGELKKALIEVLOPLAEHQARKREVDLEYKEMPRKLSF 421
Db 413 DDDKLEQIRKDYSGAMLTGELKKALIEVLOPLAEHQARKREVDLEYKEMPRKLSF 472
Oy 422 DFQ 424
Db 473 HYQ 475

```

RESULT 4  
SYM\_MOUSE STANDARD: PRT: 481 AA.

```

AC P32921:
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA ligase)
DE (TrpRS).
GN WARS OR WRS.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95018226; PubMed=7932716;
RA Rajot B., Sarger C., Bonnet J., Garret M.;
RT "An alternative splicing modifies the C-terminal end of tryptophanyl-
tRNA synthetase in murine embryonic stem cells.";
RL J. Mol. Biol. 242:599-603(1994).
CC -I- CATALYTIC ACTIVITY: AMP + L-tryptophan + tRNA(Trp) = AMP +
diphosphate + L-tryptophanyl-tRNA(Trp).
CC -I- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -I- ALTERNATIVE PRODUCTS: TWO FORMS OF THIS PROTEIN ARE PRODUCED BY
ALTERNATIVE SPLICING. A ISOFORM OF 475 RESIDUES, FOUND IN MOST
TISSUES AND A C-TERMINALLY EXTENDED FORM OF 481 RESIDUES FOUND IN
EMBRYONIC STEM CELLS.
CC -I- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
CC -I- SIMILARITY: CONTAINS 1 "WHEP-TRS" DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X69656; CAA49347.1; -;
DR EMBL; X69657; CAA49348.1; -;
DR PIR; S31461; S31461.
DR MGI; MGI:104630; Wars.
DR InterPro; IPR000738; WHEP-TRS.
DR InterPro; IPR002305; tRNA-synt_1b.
DR InterPro; IPR001412; tRNA-synt_1.
DR InterPro; IPR002306; tRNA-synt_1tp.
DR Pfam; PF00579; tRNA-synt_1b; 1.

```

```

DR Pfam; PF00458; WHEP-TRS; 1.
DR PRINTS; PS001039; TRNASYNTTRP.
DR PROSITE; PS00178; AA tRNA_LIGASE_1; 1.
DR PROSITE; PS00762; WHEP-TRS; 1.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; AMP-binding;
KW Alternative splicing.
FT DOMAIN 23
FT SITE 168 177 WHEP-TRS.
FT SITE 353 357 "HIGH" REGION.
FT VARSPLIC 476 481 "KMSKS" REGION.
SO SEQUENCE 481 AA; 54282 MW; B05A452C08074F52 CRC64;
Query Match 90.5%; Score 2032; DB 1; Length 481;
Best local similarity 89.6%; Pred. No. 2,8e-161;
Matches 380; Conservative 22; Mismatches 22; Indels 0; Gaps 0;

```

```

Oy 1 MSYKAGEDYKADCPGPNPAPTSNHGPDATBAEDVDVPMVTQTSKAGIDYKLYRF 60
Db 52 MSYKAMGEYKAGCPGPNPAGRNCDSDATKASDFVDPMTVRTSASKGIDYKLYQP 111
Oy 61 GSKIDKELINRERATGQRPHFLRGIFFSHRDMNOVLAYENKKPFYLYTGRPSSE 120
Db 112 GSKIDKELINRERATGQRPHFLRGIFFSHRDMNOVLAYENKKPFYLYTGRPSSE 171
Oy 121 AMHGHILPFIETKWLQADVFNVPVLYQMTDDEKYLKMDLTLDQAYGDAVENAKDIACGF 180
Db 172 AMHGHILPFIETKWLQADVFNVPVLYQMSDDEKYLKMDLTLDQAYGDAVENAKDIACGF 231
Oy 181 DINKTFIFSDLDYMGMSGSGFKYKVVVKIOKHVFNQVKGIFGFDSDCICKISPPAIQAAP 240
Db 232 DINKTFIFSDLDYMGMSGSGFKYKVVVKIOKHVFNQVKGIFGFDSDCICKISPPAIQAAP 291
Oy 241 SFSNFPFOIFRDRDTIOCLIPCAIDODPYFRMRDVAAPRIGYKPPALHSTFFPALQAQ 300
Db 292 SFSNFPFOIFRDRDTIOCLIPCAIDODPYFRMRDVAAPRIGYKPPALHSTFFPALQAQ 351
Oy 301 TKMSADPNSSIFLDTAKQIKTKVKNKHAFFSGGRDTIEEHROFGGNCDDVDSFMYLTFE 360
Db 352 TKMSADPNSSIFLDTAKQIKTKVKNKHAFFSGGRDTIEEHROFGGNCDDVDSFMYLTFE 411
Oy 361 EDDKLEQIRKDYSGAMLTGELKKALIEVLOPLAEHQARKREVDLEYKEMPRKLS 420
Db 412 EDDKLEQIRKDYSGAMLTGELKKALIEVLOPLAEHQARKREVDLEYKEMPRKLS 471
Oy 421 DFQ 424
Db 472 HFQ 475

```

RESULT 5  
SYM\_SCHPO STANDARD: PRT: 395 AA.

```

AC Q09692:
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Putative tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA
ligase) (TrpRS).
GN SPAC2F.13C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Gentles S., Churcher C.M., Barrell B.G., Rajandream M.A., Walsh S.V.;
RL Submitted (JUL-1995) to the EMBL/Genbank/DBJ databases.
CC -I- CATALYTIC ACTIVITY: AMP + L-tryptophan + tRNA(Trp) = AMP +
diphosphate + L-tryptophanyl-tRNA(Trp).
CC -I- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
CC -----

```

```

CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
CC -1- SIMILARITY: CONTAINS 1 "WHEP-TRS" DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X53918; CAA37872.1; -.
CC EMBL: X52113; CAA36356.1; -.
CC EMBL: M74074; AAA30799.1; -.
CC PIR: A40279; YMR0.
CC PTR: S14540; S14540.
CC InterPro: IPR000738; WHEP-TRS.
CC InterPro: IPR002305; tRNA-synt_1b.
CC InterPro: IPR001412; tRNA-synt_1.
CC InterPro: IPR002306; tRNA-synt_trp.
CC Pfam: PF00579; tRNA-synt_1b; 1.
CC Pfam: PF00458; WHEP-TRS; 1.
CC PRINTS: PR01039; TRNASYNTHTRP.
CC PROSITE: PS00178; AA-TRNA_LIGASE_I; 1.
CC PROSITE: PS00762; WHEP-TRS; 1.
CC Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding.
CC FT DOMAIN 24 69 WHEP-TRS.
CC FT SITE 1 117 DISPENSABLE TO THE CATALYTIC ACTIVITY.
CC FT SITE 169 178 "HIGH" REGION.
CC FT SITE 353 357 "KMSKS" REGION.
CC FT CONFLICT 17 17 L -> M (IN REF. 2).
CC SQ SEQUENCE 475 AA; 53729 MW; F7E531750137EB32 CRC64;

```

```

Query Match          95.0%; Score 2134.5; DB 1; Length 475;
Best Local Similarity 95.0%; Pred. No. 8,4e-170;
Matches 402; Conservative 9; Mismatches 11; Indels 1; Gaps 1;

```

```

QY 2 SYKAAAGEDYKADCPGNPAFTSNHGPDATAEEDFVDPMTVQTSAAKGIDYDKLYRFG 61
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 54 SYKAAATEDYKADCPGNPAFTSNHGPDATAEEDFVDPMTVQTSAAKGIDYDKLYRFG 113

QY 62 SSKIDKELINRIERATGQRPHFLRGIFFSHRDMNOVLAYENKKPFYLYTGRGPSSSEA 121
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 114 SSKIDKELVNRIERATGQRPHFLRGIFFSHRDMNOVLAYENKKPFYLYTGRGPSSSEA 173

QY 122 MHVGHLLPFTFTKWLQOVFNPLVIQMTDDEKYLKMDLTLDQAVGDAVENAKDIIACGFD 181
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 174 MHVGHLLPFTFTKWLQOVFNPLVIQMTDDEKYLKMDLTLDQAVGDAVENAKD-ITCGFD 232

QY 182 INKTFIFSDLDYMGSSGFGYKNVYKIOKHVTFNOVKGIFGFTSDCIGKISFPALIOAPS 241
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 233 INKTFIFSDLDYMGSSGFGYKNVYKIOKHVTFNOVKGIFGFTSDCIGKISFPALIOAPS 292

QY 242 FSNSEPOIFRDRDIOCLIPCALIDODPEYFRMTRDVAPRIGYKPKALLHSFFPALQAGQT 301
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 293 FSNSEPOIFRDRDIOCLIPCALIDODPEYFRMTRDVAPRIGYKPKALLHSFFPALQAGQT 352

QY 302 KMSASDPNNSIFLDTAKQIKTKYKNAHAFSGRPTIEHQFGNCVNDVSFMTLTFPLE 361
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 353 KMSASDPNNSIFLDTAKQIKTKYKNAHAFSGRPTIEHQFGNCVNDVSFMTLTFPLE 412

QY 362 DDDLEQIRKDYTSGAMLTGELKALIEVLOPLIAEHQARKEVTDIVKEFMTPRKLSF 421
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 413 DDDLEQIRKDYTSGAMLTGELKALIEVLOPLIAEHQARKEVTDIVKEFMTPRKLSY 472

QY 422 DFQ 424
   |||
Db 473 DFQ 475

```

```

RESULT 3
SYM_RABIT
ID SYM_RABIT          STANDARD;          PRT;          475 AA.

```

```

AC P23612; Q28607;
DT 01-NOV-1991 (Rel. 20, Created)
DT 15-JUN-1998 (Rel. 36, Last sequence update)
DT 15-JUN-1998 (Rel. 36, Last annotation update)
DE tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA ligase)
DE (trpRS).
GN WARS.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90239043; PubMed=2185472;
RA Lee C.C., Craigen W.J., Muzny D.M., Harlow E., Caskey C.T.;
RT "Cloning and expression of a mammalian peptide chain release factor
RT with sequence similarity to tryptophanyl-tRNA synthetases.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:3508-3512(1990).
RN [2]
RP REVISIONS TO 169-174 AND 227-228, AND FUNCTION.
RX MEDLINE=94009008; PubMed=8404867;
RA Frolova L.Y., Dalphin M.E., Justesen J., Powell R.J., Druegon G.,
RA McCaughan K.K., Kisselev U.L., Tate W.P., Haenni A.-L.;
RT "Mammalian polypeptide chain release factor and tryptophanyl-tRNA
RT synthetase are distinct proteins.";
RL EMBO J. 12:4013-4019(1993).
CC -1- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +
CC diphosphate + L-tryptophanyl-tRNA(Trp).
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
CC -1- SIMILARITY: CONTAINS 1 "WHEP-TRS" DOMAIN.
CC -1- CAUTION: WAS ORIGINALLY (REF.1) THOUGHT TO BE AN EUKARYOTIC
CC RELEASE FACTOR (ERF).

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```

```

CC EMBL: M33460; AAA31246.1; ALT_SEQ.
CC EMBL: U02595; AAB60257.1; -.
CC PIR: A35904; YMRBPR.
CC InterPro: IPR000738; WHEP-TRS.
CC InterPro: IPR002305; tRNA-synt_1b.
CC InterPro: IPR001412; tRNA-synt_1.
CC InterPro: IPR002306; tRNA-synt_trp.
CC Pfam: PF00579; tRNA-synt_1b; 1.
CC Pfam: PF00458; WHEP-TRS; 1.
CC PRINTS: PR01039; TRNASYNTHTRP.
CC PROSITE: PS00178; AA-TRNA_LIGASE_I; 1.
CC PROSITE: PS00762; WHEP-TRS; 1.
CC Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding.
CC FT DOMAIN 23 68 WHEP-TRS.
CC FT SITE 168 177 "KMSKS" REGION.
CC FT SITE 353 357 "HIGH" REGION.
CC SQ SEQUENCE 475 AA; 53799 MW; 33BC9E718FF43DC4 CRC64;

```

```

Query Match          91.5%; Score 2056; DB 1; Length 475;
Best Local Similarity 90.5%; Pred. No. 2,8e-163;
Matches 383; Conservative 20; Mismatches 20; Indels 0; Gaps 0;

```

```

QY 2 SYKAAAGEDYKADCPGNPAFTSNHGPDATAEEDFVDPMTVQTSAAKGIDYDKLYRFG 61
   |||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 53 SYKAAAGEDYKADCPGNPSHPDSEADVDKEDFVDPMTVQTSAAKGIDYDKLYRFG 112

QY 62 SSKIDKELINRIERATGQRPHFLRGIFFSHRDMNOVLAYENKKPFYLYTGRGPSSSEA 121
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 113 SSKIDKELVNRIERATGQRPHFLRGIFFSHRDMNOVLAYENKKPFYLYTGRGPSSSEA 172

QY 122 MHVGHLLPFTFTKWLQOVFNPLVIQMTDDEKYLKMDLTLDQAVGDAVENAKDIIACGFD 181

```

RA Erolova L.Y., Grigorieva A.Y., Sudomolina M.A., Kisselev L.L.;  
 RT "The human gene encoding tryptophanyl-tRNA synthetase: interferon-  
 response elements and exon-intron organization.";  
 RL Gene 128:237-245(1993).  
 [7]  
 RP SEQUENCE OF 265-276; 278-296; 299-317 AND 350-365.  
 RC TISSUE=Keratinocytes;  
 RX MEDLINE=93162043; PubMed=1286667;  
 RA Raamussen H.H., van Damme J., Puype M., Geesser B., Celis J.E.,  
 RT Vanderkerckhove J.;  
 "Microsequences of 145 proteins recorded in the two-dimensional gel  
 protein database of normal human epidermal keratinocytes.";  
 RL Electrophoresis 13:960-969(1992).  
 RN  
 RP  
 RX MEDLINE=9225128; PubMed=1373391;  
 RA Bange F.-C., Flohr T., Buwitt U., Boettger E.C.;  
 RT "An interferon-induced protein with release factor activity is a  
 tryptophanyl-tRNA synthetase.";  
 RL FEBS Lett. 300:162-166(1992).  
 CC -1- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +  
 CC diphosphate + L-tryptophanyl-tRNA(Trp).  
 CC -1- SUBUNIT: HOMODIMER.  
 CC -1- INDUCTION: BY INTERFERON GAMMA.  
 CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.  
 CC -1- SIMILARITY: CONTAINS 1 "WHEP-TRS" DOMAIN.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: M77804; AAA67324.1; -  
 DR EMBL: X59892; CAA42545.1; -  
 DR EMBL: M61715; AAA61298.1; -  
 DR EMBL: X62570; CAA44450.1; -  
 DR EMBL: S82905; AAB39381.1; -  
 DR EMBL: X67920; CAB94198.1; -  
 DR EMBL: X67921; CAB94198.1; JOINED.  
 DR EMBL: X67922; CAB94198.1; JOINED.  
 DR EMBL: X67923; CAB94199.1; -  
 DR EMBL: X67924; CAB94199.1; JOINED.  
 DR EMBL: X67925; CAB94199.1; JOINED.  
 DR EMBL: X67926; CAB94199.1; JOINED.  
 DR EMBL: X67927; CAB94199.1; JOINED.  
 DR EMBL: X67928; CAB94199.1; JOINED.  
 DR PIR: A41706; A41706.  
 DR PIR: A41633; A41633.  
 DR PIR: JH0533; JH0533.  
 DR PIR: S19246; S19246.  
 DR Aarhus/Chent-2DPAGE; 3524; IEF.  
 DR PDBT-2DPAGE; P23381; -  
 DR MIM: 191050; -  
 DR InterPro: IPR000738; WHEP-TRS.  
 DR InterPro: IPR002305; tRNA-synt\_1b.  
 DR InterPro: IPR001412; tRNA-synt\_1.  
 DR InterPro: IPR002306; tRNA-synt\_1tp.  
 DR Pfam: PF00579; tRNA-synt\_1b; 1.  
 DR Pfam: PF00458; WHEP-TRS; 1.  
 DR PRINTS: PRO1039; TRNASYNTHTRP.  
 DR PROSITE: PS00178; AA\_TRNA\_LIGASE\_1; 1.  
 DR PROSITE: PS00762; WHEP-TRS; 1.  
 KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding.  
 FT DOMAIN 19 64 WHEP-TRS.  
 FT SITE 164 173 "HIGH" REGION.  
 FT SITE 349 353 "KMSKS" REGION.  
 FT CONFLICT 213 214 SY -> GD (IN REF. 3).  
 FT CONFLICT 424 424 A -> R (IN REF. 4).  
 SQ SEQUENCE 471 AA; 53165 MW; E9634444905A0D0 CRC64;

Query Match 99.3%; Score 2231; DB 1; Length 471;  
 Best Local Similarity 99.5%; Pred. No. 8e-178;  
 Matches 422; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 MSYKAAGEDYKADCPGCPNPAPTSNHEGDATEAEEDVDPMTVOTSSAKGIDYOKLYRF 60  
 DB 48 MSYKAAGEDYKADCPGCPNPAPTSNHEGDATEAEEDVDPMTVOTSSAKGIDYOKLYRF 107  
 QY 61 GSSKIDKELIRIRATGQRPHHLLRGIFPSHRDMOVLDAYENKKKPFYLYTGRGSSSE 120  
 DB 108 GSSKIDKELIRIRATGQRPHHLLRGIFPSHRDMOVLDAYENKKKPFYLYTGRGSSSE 167  
 QY 121 AMHGHILPFIETKWLQDFVNPPLVIOQMTDDEKYLKMDLTLDQAYGDAVENAKDIIACGF 180  
 DB 168 AMHGHILPFIETKWLQDFVNPPLVIOQMTDDEKYLKMDLTLDQAYGDAVENAKDIIACGF 227  
 QY 181 DINKTFISDLIDYMGSSGFKNVYKIOKHYTPNOVKCIFSTSDICGKISPAIQAP 240  
 DB 228 DINKTFISDLIDYMGSSGFKNVYKIOKHYTPNOVKCIFSTSDICGKISPAIQAP 287  
 QY 241 SESNSFQIFRDRIIDICLIPCAIDODPYFPMTRDVAAPRIGYKPKALLHSFFPALOGAO 300  
 DB 288 SFSNSFQIFRDRIIDICLIPCAIDODPYFPMTRDVAAPRIGYKPKALLHSFFPALOGAO 347  
 QY 301 TKMSASDPNSSIFLDTAKQIKTKVKNKHAFSGGRDTEENRQFGNCDDVDSFMYLTFFL 360  
 DB 348 TKMSASDPNSSIFLDTAKQIKTKVKNKHAFSGGRDTEENRQFGNCDDVDSFMYLTFFL 407  
 QY 361 EDDDKLEIRKDYISGAMLTGELKALIEVLOPLAEHQARKREYTDIYKEFMTPKLS 420  
 DB 408 EDDDKLEIRKDYISGAMLTGELKALIEVLOPLAEHQARKREYTDIYKEFMTPKLS 467  
 QY 421 FDFQ 424  
 DB 468 FDFQ 471  
 RESULT 2  
 SYW\_BOVIN  
 ID SYW\_BOVIN STANDARD; PRT; 475 AA.  
 AC P17248;  
 DT 01-AUG-1990 (Rel. 15, Created)  
 DT 01-MAY-1992 (Rel. 22, Last sequence update)  
 DT 01-FEB-1996 (Rel. 33, Last annotation update)  
 DE Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA ligase)  
 GN WARS.  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Bovinae; Bos.  
 OX NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.  
 RC TISSUE=Retina;  
 RX MEDLINE=91329348; PubMed=1907847;  
 RA Garret M., Pajot B., Trezeguet V., Labouesse J., Merle M.,  
 RA Gandar J.-C., Benedetto J.-P., Sallafrance M.-L., Alterio J.,  
 RA Gueguen M., Sarger C., Labouesse B., Bonnet J.;  
 "A mammalian tryptophanyl-tRNA synthetase shows little homology to  
 RT prokaryotic synthetases but near identity with mammalian peptide  
 RT chain release factor.";  
 RL Biochemistry 30:7809-7817(1991).  
 RP  
 RP SEQUENCE OF 17-475 FROM N.A.  
 RC TISSUE=Pancreas;  
 RA Garret M., Trezeguet V., Pajot B., Gandar J.-C., Merle M.,  
 RA Gueguen M., Benedetto J.-P., Sarger C., Alterio J., la Bouessec B.,  
 RA Labouesse J., Bonnet J.;  
 CC Submitted (MAR-1990) to the EMBL/GenBank/DBJ databases.  
 CC -1- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +  
 CC diphosphate + L-tryptophanyl-tRNA(Trp).  
 CC -1- SUBUNIT: HOMODIMER.

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 24, 2002, 12:49:40 : Search time 7.77387 Seconds  
(without alignments)  
2111.829 Million cell updates/sec

Title: us-09-813-718-10\_COPY\_48\_471

Perfect score: 2246  
Sequence: 1 MSYKAAAGEDYKADCPGNP.....VTDEIVKEMTPRKLSFDFQ 424

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2231	99.3	471	1	SYW_HUMAN
2	2134.5	95.0	475	1	SYW_BOVIN
3	2056	91.5	475	1	SYW_RABIT
4	2032	90.5	481	1	SYW_MOUSE
5	1210	53.9	395	1	SYW_SCHRO
6	1163	51.8	432	1	SYW_YEAST
7	803	35.8	385	1	SYW_PYRAB
8	409.5	18.2	370	1	SYW_METJA
9	397.5	17.7	364	1	SYW_METJA
10	386	17.2	134	1	SYW_ENCCU
11	370.5	16.5	420	1	SYW_ARCPU
12	353	15.7	374	1	SYW_AERPE
13	192	8.5	323	1	SYW_ARCPU
14	175.5	7.8	341	1	SYW_CLOLO
15	172	7.7	394	1	SYW_YEAST
16	170.5	7.6	366	1	SYW_SULSO
17	169	7.5	395	1	SYW_AQUAE
18	166.5	7.4	346	1	SYW_CHLTR
19	163.5	7.3	353	1	SYW_BORBU
20	159.5	7.1	337	1	SYW_TREPA
21	158	7.0	346	1	SYW_CHLMU
22	152	6.8	344	1	SYW_CHLPN
23	146	6.5	326	1	SYW_HELPY
24	138	6.1	319	1	SYW_METH
25	136.5	6.1	528	1	SYW_BOVIN
26	135.5	6.0	328	1	SYW_THEMA
27	133	5.9	328	1	SYW_BACST
28	132.5	5.9	343	1	SYW_MYCLE
29	132.5	5.9	528	1	SYW_HUMAN
30	130	5.8	326	1	SYW_HELPJ
31	129	5.7	401	1	SYW_SCHPO
32	128.5	5.7	379	1	SYW_YEAST
33	128	5.7	334	1	SYW_HAEIN

## ALIGNMENTS

RESULT 1	SYW_HUMAN	STANDARD:	PRT:	471 AA.
AC	P23381; P78535; Q9UDL3;			
DT	01-NOV-1991 (Rel. 20, Created)			
DT	01-DEC-1992 (Rel. 24, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA ligase)			
DE	(TrpRS) (IIP53) (hmRS).			
GN	WARS.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
OX	NCBI_Taxid=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=92105071; Pubmed=1761529;			
RA	Rubin B.Y., Anderson S.L., Xing L., Powell R.J., Tate W.P.;			
RT	"Interferon induces tryptophanyl-tRNA synthetase expression in human			
RT	fibroblasts.";			
RL	J. Biol. Chem. 266:24245-24248(1991).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=92107982; Pubmed=1763065;			
RA	Flecker J., Rasmussen H.H., Justesen J.;			
RT	"Human interferon gamma potentially induces the synthesis of a 55-kDa			
RT	protein (gamma 2) highly homologous to rabbit peptide chain release			
RT	factor and bovine tryptophanyl-tRNA synthetase.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 88:11520-11524(1991).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=92112058; Pubmed=1765274;			
RA	Prolova L.Y., Sudomolina M.A., Grigor'eva A.Y., Zinov'eva O.L.,			
RT	Kisselev L.L.;			
RT	"Cloning and nucleotide sequence of the structural gene encoding for			
RT	human tryptophanyl-tRNA synthetase.";			
RL	Gene 109:291-296(1991).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=92164636; Pubmed=1537332;			
RA	Buwaitt U., Flohr T., Boettger E.C.;			
RT	"Molecular cloning and characterization of an interferon induced			
RT	human cDNA with sequence homology to a mammalian peptide chain			
RT	release factor.";			
RL	EMBO J. 11:489-496(1992).			
RN	[5]			
RP	SEQUENCE OF 1-13 FROM N.A.			
RX	MEDLINE=96319994; Pubmed=8724762;			
RA	Sokolova I.V., Narovilianskii A.N., Amchenkova A.M., Turpaev K.T.;			
RT	"Alternative splicing of 5'-terminal exons of the human tryptophanyl-			
RT	tRNA synthetase gene.";			
RL	Mol. Biol. (Mosk) 30:319-329(1996).			
RN	[6]			
RP	SEQUENCE OF 1-141 AND 182-471 FROM N.A.			
RC	TISSUE-Sperm;			
RX	MEDLINE=93292992; Pubmed=7685728;			

34	124.5	5.5	347	1	SYW_MYCCE	P47372 mycoplasma
35	122	5.4	350	1	SYW_CABEL	P46579 caenorhabdit
36	118.5	5.3	343	1	SYW_CIOAB	O97166 clostridium
37	116.5	5.2	330	1	SYW_BACSU	P21656 bacillus su
38	112.5	5.0	335	1	SYW_BUCAI	P57602 buchnera ap
39	110	4.9	306	1	SYW_METJA	O57834 methanococ
40	109.5	4.9	336	1	SYW_MYCRU	O53386 mycobacteri
41	107	4.8	330	1	SYW_RICPR	O92d76 rickettsia
42	106.5	4.7	334	1	SYW_ECOLI	P00954 escherichia
43	105.5	4.7	333	1	SYW_PASMO	P57956 pasteurella
44	105	4.7	426	1	SYW_MYCLE	O49900 mycobacteri
45	104	4.6	337	1	SYW_SYNY3	P73655 synechocyst



DB 342 IAMEKLERFLAEHOSRLEKAKTIANKLVEPPR 373

Search completed: October 24, 2002, 12:53:53  
Job time : 16.2521 secs

Db 331 ECKNNTAEFIKRFEEELSVKREK 353

RESULT 13

T43806

Cryptophan--tRNA ligase (EC 6.1.1.2) [imported] - Encephalitozoon cuniculi (fragment)

C:Species: Encephalitozoon cuniculi

C:Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 02-Sep-2000

C:Accession: T43806

R:Peptrellallade, E.; Broussolle, V.; Peyret, P.; Metenier, G.; Gouy, M.; Vlvares, C.P. Mol. Biol. Evol. 15, 683-689, 1998

A:Title: Microsporidia, amitochondrial protists, possess a 70-kDa heat shock protein gene

A:Reference number: 222693; MUID:98277683

A:Accession: T43806

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Residue type: DNA

A:Molecule type: DNA

A:Residues: 1-134 <PE>

C:Cross-references: EMBL:AJ012470; PIDN:CAA10034.1

C:Superfamily: mammalian tryptophan--tRNA ligase; amino acid--tRNA ligase repeat homolog

C:Keywords: ligase

[illegible]

**RESULT**   **14**

E69461  
cryptophanyl-LRNA synthetase (trps) homolog - Archaeoglobus fulgidus  
**C:**Species: Archaeoglobus fulgidus  
**C:**Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 21-Jul-2000  
**O:**Accession: E69461  
**R:**Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, T.E.; Nelson, K.E.; Ketchum, K.A.; Dodson,  
. ; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirness, E.F.  
; Glöckner, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.

Nature 390, 364-370, 1997

A:Authors : utterback, T.: Cotton, M.D.: Spriggs, T.: Artlich, P.: Kalne, B.P.: Sykes, S.  
Smith, H.O.: Woese, C.R.: Venter, J.C.

A>Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeobacterium Pyrococcus furiosus strain ATCC 49239

A:Title number: A69250; WUID:98049343

A:Accession : E69461

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residue: 1..420 <**KLEP**>

A:Cross-references: GB:AEO00986; GB:AEO00782; NTD:g2689309; PIDN:AAE89554.1; PID:g264885

A:Superfamily: mammalian tryptophan--tRNA ligase; amino acid--tRNA ligase repeat homology

```

Query Match      16.5%; Score 370.5; DB 2; Length 420;
Best Local Similarity 27.3%; Pred. No. 2,9e-22;
Matches 118; Conservative 60; Mismatches 166; Indels 89; Gaps

QY      38 VDPPTVQTSSAKGIDYDKLIYRFGSSKIDKMLIRERATGQRPHHFLRGISFSSHDMN 97
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db       3 VTPEVEGV-----IDSKLIEFGMOPE-SEVLEPID-----NPHILMRKQATFGHRDYW 52

QY      98 QVDAAYENKKRFYLYLTGSGSSSEAMHGHILPFTFKMDVENVPLVYQMTDEKTLMK 157
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db       53 RILEAMQKKEPMAWMSGFMSG-LPHFGKHKTMDIELIWHOSAGSKAEV-AIADAEASHVR 110

QY      158 DLTLDQAYGADEVNAKDIACGFPIKNTFIISDLDYMGSSGFKNV-KIQKHVTPTNOV 216
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

D	b	111	GLSMKRTRELGMXIKSLNALGLREDAVITYQS-----KSNHKDLAFELSAENVNSSEL	164
Q	y	217	KGIFGFTDSDCIGKISFPALDAPFSNSPQIFRDRDTDIOCLPCAIODDPYFRMTRDY	276
D	b	165	RAIFGFSNDSTLAKMFVTAIOADIL--HPOLSDFGPKPVVYPVAGADODPHMRLTRDL	221
Q	y	277	APRI-----	280
D	b	222	AARISIFSEEPVEGGVNRSHRSKKAENVLSLRLDLEFDKKIYEHMDIFGEAEIETRAVRKI	281
Q	y	281	-----GYPRALLHSTFFPALOGAOTKKASASPNSIFELDTAKOIKTKVKNKAFSGG	333
D	b	282	EVEIGGEAFIPSSYTHRFHTGLTGG--KMSSSKPESYISLLOPPEGARKVVK--ATGG	338
Q	y	334	RDTEEHROFGNCDVNVSPMYLTFPLED--DDKLEQIRKYDTSGAMLTGELKALALEVLO	392
D	b	339	RATAEQGRLOGGEDRCVFEFELYSFHLIDSEBLNTEACBREGRRLLCGCRKMAAEVLYK	398
Q	y	393	PLIAEHQARKEV	405
D	b	399	SFLKEHDEKMEAV	411

RESULT 15  
D72477  
probable tryptophanyl-tRNA synthetase APE2461 - Aeropyrum pernix (strain K1)  
C:Species: Aeropyrum pernix  
C:Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 20-Jun-2000  
C:Accession: D72477  
R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Tawa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudo, Y.; Yamazaki, J.  
DNA Res. 6: 83-101, 1999  
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, *Aero*  
Reference number: A72450; MUID:99310339

A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-374 <KAM>  
A;Cross-references: DDBJ:AP000064; NID:95105945; PIDN:BAA81476.1; PID:95106165  
A;Experimental source: strain K1  
C;Genetics:  
A;Gene: APE2461  
A;Superfamily: mammalian tryptophan--tRNA ligase; amino acid--tRNA ligase repeat homo

Query Match	15.7%	Score 353;	DB 2;	Length 374;
Best Local Similarity	30.4%	Pred. No. 6.5e-21;		
Matches 119;	Conservative 70;	Mismatches 165;	Indels 38;	Gaps 17/;
QY	38	VDPWTVQTSAAKGIYDYLIVFGSSKIDKELINRTERATGGRPHNLFGRGIFGSHRDN	97	
	8	LDPM-----GAAVEIKDYDLRTFTGGIRPSEVL--PLRKACNRPSEFLMRGIIIFGRHED	61	
Db				
QY	98	QVLDAENKKRFFYLITGRGPSSSEAHNVHLIRPFTKWLQ--DVFNVPRLVLTQMTDEKYLIM	156	
	62	KILEKAKGERVAALVLTGMPGSK--FHGHNKLTVDLITLYQKNGFV--FVALADAEAPAV	118	
Db				
QY	157	KDLTLDQAVDAVEN--AKDIIACGFDIKNT--FIFSDDLVMGMSGFGYKNVVKIQKHVFEN	214	
	119	RRIGEEAVRILAVEEYLIANMIALGSDPRDTEVFQ----TNGTGYEFLRIQLGFSKRVITA	174	
Db				
QY	215	QVKGIYFG--FTTSDSICIGKISFPRAIQAPSFNSPFOIFRDRITDIOCLIPCAIDODDEYFRMT	273	
	175	EMEAIVYGLGLTAKMMASISL-----QAADLIHLQDLDEYGGV--HYVVVGAQDDPHLRILT	227	
Db				
QY	274	RDVAVR---IGYRPALLHSTFEPALQGAQTQMSASDPNSSIFLDTAKQIKTKVKNKA	329	
	228	RDLADRMAGVLELRPASTYHKLQGLGD--RKMSSSRPDSITFLDPREVAARLFR-A	284	
Db				
QY	330	FSGGRTTEEHRQEGNCQV--DVSEFWYLTFFLELDDDKLEQIRKDYTS---GAMLTGELK	385	
	285	LTGGRATAEQGRILGIVPEVCSVYHMDLYHLMPRDGVEKHI---YTSCHLGIKILGECQO	341	
Db				
QY	386	ALIEVLOPLIAEHQARRKEVTDIEYKERTMPTR	417	

[illegible]

```

RESULT      11
Fc6476     tryptophan--tRNA ligase (EC 6.1.1.2) - Methanococcus jannaschii
N;Alternate names: tryptophanyl-tRNA synthetase
C;Species: Methanococcus jannaschii
C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jul-2000
Accession: Fc6476
R;Butl, C.J.; White, O.; Olsen, G.-J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake, J.; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodok, A.; Tsom, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A. Science 273, 1058-1073, 1996
A;Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii
A;Reference number: A64300; MUID:96337999
A;Accession: Fc6476
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A;Residues: 1-370 <BDU>
A:Cross-references: GB:U67582; GB:L77117; NID:g1592064; PID:AAB99425.1; PID:g1592065; TCGeneatics:
A;Map position: FOR1375885-1376997
A;Start codon: GTG
C;Superfamily: mammalian tryptophan--tRNA ligase; amino acid--tRNA ligase repeat homolog
C;Keywords: aminoaacyl-tRNA synthetase; ligase; protein biosynthesis

Query Match          18.2%; Score 409.5; DB 2; Length 370;
Best Local Similarity 30.9%; Pred. No. 1.6e-25;
Matches 119; Conservative 72; Mismatches 153; Indels 41; Gaps 16;

Dy       40 PTWTSSAKGDYDKLVFGSSKIDKELNRIERATGRCPRHFLRGIFFSHRDMQV 99
           ||| :||| |::| | |   | | | | | | | | | | | | | | | |
Db        8 PW-ETPAV--IDRYKTMDGFVKRPVDGLDAKE-----HHFRKNIIILGHDPERI 57

Oy      100 LDAYENKKPFYLTYTGRGPSEAMHYGHLIPFIPTKWLD--DVFNVPRLVIOMTDDEKYI 156
           |:| |::| | | | | | | | | | | | | | | | | | | | | | | |
Db       58 VDAIKNNKFVAIVSGAMPSPSK-MHFGHKMNVDDLRFYGKYNINNIPE----ADLEAVWA 112

Oy     157 KDLLDAQAGAV-ENAKDIACGFDINKTFPSDLDTMGSSGEFYKANVMI-QKHVTEN 214
           ::::| | | | | | | | | | | | | | | | | | | | | | | |
Db    113 RMSETTKELALNTNYINYIALGIAPPERINVLYLSOKROY---KDALLSLSRFTMS 167
```

Qy	215	QVKGFGTDBDCIGKISFPAIQAAPSFSNSFPQIFRPT--DIQCLIPCAIDDDPPYFRM	272
Db	168	EMKAIYGGKGTNGIHVAPAIYQVADLL--HPQIDENLSEPKRVVYPGVIDDPIHRL	224
Qy	273	TRDVAPR--IGYKPAALHSTFFPALOGAQTKASASPNSIFLTPAKOIKTKVNKA	329
Db	225	TRDIANRAKKEKFIIPSSYTHRFMTGLLG--KKSSKSPETALITLTDDEKIVKKRITS-A	281
Qy	330	FSGGHDITIEHRQFG--NCDVDVSEWYTLFFLEDDKLQAIKKDYTSGLMTGELKA	386
Db	282	KTGSGRETLSEHKKYGAVPEECVVELFLY--HLIIDKELAEIYOKRSGELTGCKCKM	339
Qy	387	LIEVLPFLAEHQARKREVTDEIYK	411
Db	340	AYERVEFLKDLKEKREQAKEIAVK	364

**RESULT 12**

E69131  
tryptophan--tRNA ligase (EC 6.1.1.2) - Methanobacterium thermoautotrophicum (strain D)

N:Alternate names: tryptophanyl-tRNA synthetase  
C:Species: Methanobacterium thermoautotrophicum  
C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 22-Oct-1999  
C:Accession: E69131

R:Smith, D.R.; Doucette-Stamm, L.A.; Delougnery, C.; Lee, H.; Dubois, J.; Alredge, T.  
J.; Liu, D.; Spadafora, R.; Vitcalre, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Janani,  
K.I., S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.

J: Bacterial: 179, 7135-7155, 1997  
A>Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: fu  
A:Reference number: A69000; MUID:98037514  
A:Accession: E69131  
A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-364 <MNH>  
A:Cross-references: GB:AE000812; GB:AE000666; NID:g2621298; PIDN:AAB84757.1; PID:g2622  
A:Experimental source: strain Delta H  
C:Genetics:  
A:Gene: MTH251  
A:Start codon: TTG  
C:Superfamily: mammalian tryptophan--tRNA ligase; amino acid--tRNA ligase repeat homo  
C:Keywords: aminocacyl-tRNA synthetase; ligase; protein biosynthesis

Query Match 17.7%; Score 397.5; DB 2; Length 364;  
Best Local Similarity 27.9%; Pred. NO.1.5e-24;  
Matches 107; Conservative 76; Mismatches 153; Indels 47; Gaps 14;

OY 38 VDPWTVQTSAGKSIDYDLIVRFGSSKIDKELINERATGCRPHFLRSGIFPSRDNM 97  
::: ||| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|  
DB 2 IDPW-----GSKK-LEXPDLIENFGRPF-SEVLDDV-----PEFSWMLRRIILIGHRYE 50  
  
OY 98 OVLDAVENKKRPYYLTGTSGSPSEAMNVGHILPFTFKLQDVFNVPYLVIOMTDEKYIMK 157  
::: |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|  
DB 51 RIISAKKGGEFAVVYTGMPSPGR-MHIGKHKIIVDQLRW-YPRMGALPIPIADMEAVSAR 108  
::: |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|  
  
OY 158 DLTLLQAAGDAVEN-AKDIICAGCPDK-----TFESDLDVDYGMSGGFYKNVVK 206  
::: |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|  
DB 109 GVDFEDSRIRALEEYTAGYTALGLDELKONIHVYIQSENMLWEVDLAY----- 156  
  
OY 207 IOKHTVFNQVGIGEFETSDCGIKISFPALIOAAFSNSFPQIFRDRITDLOCLPCALDQ 266  
::: |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|  
DB 157 LAGKVNFEHLRAIYGFSTGSTMAMHYAPRLIQVSDLIHQDLDELGPGR---PVIVPVRDQ 213  
  
OY 267 DPYFMRTADVARI----GYKPRLHSITFFPALDAQOTKKSASDPNSSITFLTAKOIK 322  
||: |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|  
DB 214 DPHILTLADIARFRDRYGFLLPSYSTYHRFMGGTLTG--KMSSNPKSAIFLSDTPEAE 271  
  
OY 323 TVKNHAFSGSGDTEEHROFGNCNDVVSFMYLFFLE-DDDKLEQIRKDYTGSMALTG 381  
|::: |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|  
DB 272 AKI-RNAATGTGHEITKEORELGVPFECCIIYETLLYHMSGDSRLIEITYESCRRGTLMCG 330  
  
OY 382 ELKALLIEVLOPLIAHQARRKE 404  
|::: |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|





A:Gene: WRS

A:Gene: SPDR.SPAC2F7 13C

A:Residues: 1-212, 'GD', 215-471 <FRO2>  
 A:Cross-references: GB:M61715; NID:g340367; PIDN:AAA61298.1; PID:g340368  
 A:Experimental source: fibroblast  
 C:Genetics:  
 A:Gene: GDB:WARS; IFF53  
 A:Cross-references: GDB:119632; OMIM:191050  
 A:Map position: 14q23-14q31  
 A:Introns: 33/3; 105/1; 141/2; 181/2; 242/2; 276/1; 313/3; 371/3; 418/3  
 C:Superfamily: mammalian tryptophan--tRNA ligase; amino acid--tRNA ligase repeat homolog  
 C:Keywords: aminocyl-tRNA synthetase; ATP; ligase; protein biosynthesis  
 F:19-64/Domain: amino acid--tRNA ligase repeat homology <ATL>

Query Match 99.3%; Score 2231; DB 1; Length 471;  
 Best Local Similarity 99.5%; Pred. No. 1,4e-173;

Matches 422; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY 1 MSYKAAGEDYKADCPGPNAPTSNHGPDATAEEDFVDPMTVQTSSAKGIDYDKLIVRF 60
DB 48 MSYKAAGEDYKADCPGPNAPTSNHGPDATAEEDFVDPMTVQTSSAKGIDYDKLIVRF 107
QY 61 GSKIDKELINIERATGQRPHRLRGIFFSHRDMNOVLDAVENKKPFLYLYGRGPSSA 120
DB 108 GSKIDKELINIERATGQRPHRLRGIFFSHRDMNOVLDAVENKKPFLYLYGRGPSSA 167
QY 121 AMHVGHLIPFTKWLQDVFNPLVIOMTDDEKYLKDLTDQAYGDAVENAKDIIACGF 180
DB 168 AMHVGHLIPFTKWLQDVFNPLVIOMTDDEKYLKDLTDQAYGDAVENAKDIIACGF 227
QY 181 DINKTFISDLDYMGSSGFEYKNVYKIOKHVFNQVKGIFGFTSDCIGKISFPALQAA 240
DB 228 DINKTFISDLDYMGSSGFEYKNVYKIOKHVFNQVKGIFGFTSDCIGKISFPALQAA 287
QY 241 SFSNFPQIFRRTDIOCLIPCAIDODPYFRMTDVAAPRIGYKPPALHSTFFPALQAG 300
DB 268 SFSNFPQIFRRTDIOCLIPCAIDODPYFRMTDVAAPRIGYKPPALHSTFFPALQAG 347
QY 301 TKMSADPNSSIFLDTAKQIKTKVKNKHAFFSGGRDTEEHROFGNCDDVDSFMYLTFE 360
DB 348 TKMSADPNSSIFLDTAKQIKTKVKNKHAFFSGGRDTEEHROFGNCDDVDSFMYLTFE 407
QY 361 EDDDKLEQIRKDYTGAMLTGELKALIEVLOPLIAEHQARKEVDTEYKEFMPRKLS 420
DB 408 EDDDKLEQIRKDYTGAMLTGELKALIEVLOPLIAEHQARKEVDTEYKEFMPRKLS 467
QY 421 FDFQ 424
DB 468 FDFQ 471

```

## RESULT 2

tryptophan--tRNA ligase (EC 6.1.1.2) [validated] - bovine

N:Alternate names: tryptophanyl-tRNA synthetase

C:Species: Bos primigenius taurus (cattle)

C:Date: 30-Sep-1992 #sequence\_revision 30-Sep-1992 #text\_change 21-Jul-2000

C:Accession: A40279; JN0354; S10460; S14540

R:Garret, M.; Pajot, B.; Trezeguet, V.; Labouesse, J.; Merle, M.; Gandar, J.C.; Benedetti

Biochemistry 30, 7809-7817, 1991

A:Title: A mammalian tryptophanyl-tRNA synthetase shows little homology to prokaryotic s

A:Reference number: A40279; MUID:91329348

A:Accession: A40279

A:Residues: 1-475 <GAR>

A:Cross-references: GB:M74074; EMBL:X53918; NID:g163798; PIDN:AAA30799.1; PID:g163799

A:Experimental source: pancreas

A>Note: The authors translated the codon CTG for residue 347 as Ala and CAG for residue

R:Zarganova, T.A.; Kovaleva, G.K.; Favorova, O.O.; Aevina, H.B.; Telezhinskaya, I.H.

Bioorg. Khim. 15, 1307-1311, 1989

A:Title: Amino acid sequence of several peptides of tryptophanyl-tRNA synthetase from ce

A:Reference number: JN0354; MUID:90211408

A:Accession: JN0354

A:Molecule type: protein

A:Residues: 112-124;282-287, 'N',288, 'F',289-292, 'Q',293-294, 'R',336-353;423-441,443-  
 A:Experimental source: liver  
 A>Note: this paper is in Russian  
 C:Superfamily: mammalian tryptophan--tRNA ligase; amino acid--tRNA ligase repeat homo  
 C:Keywords: aminocyl-tRNA synthetase; ATP; ligase; protein biosynthesis  
 F:24-69/Domain: amino acid--tRNA ligase repeat homology <ATL>

Query Match 95.0%; Score 2134.5; DB 1; Length 475;  
 Best Local Similarity 95.0%; Pred. No. 1e-165;

Matches 402; Conservative 9; Mismatches 11; Indels 1; Gaps 1;

```

QY 2 SYKAAGEDYKADCPGPNAPTSNHGPDATAEEDFVDPMTVQTSSAKGIDYDKLIVRF 61
DB 54 SYKATGEDYKVDCEPGDAPESGEGDLATAEDEDFVDPMTVQTSSAKGIDYDKLIVRF 113
QY 62 SSKIDKELINIERATGQRPHRLRGIFFSHRDMNOVLDAVENKKPFLYLYGRGPSSA 121
DB 114 SSKIDKELINIERATGQRPHRLRGIFFSHRDMNOVLDAVENKKPFLYLYGRGPSSA 173
QY 122 MHVGHILIPFTKWLQDVFNPLVIOMTDDEKYLKDLTDQAYGDAVENAKDIIACGF 181
DB 174 MHVGHILIPFTKWLQDVFNPLVIOMTDDEKYLKDLTDQAYGDAVENAKDIIACGF 232
QY 182 INKTFISDLDYMGSSGFEYKNVYKIOKHVFNQVKGIFGFTSDCIGKISFPALQAA 241
DB 223 INKTFISDLDYMGSSGFEYKNVYKIOKHVFNQVKGIFGFTSDCIGKISFPALQAA 292
QY 242 FSNSEFPQIFRRTDIOCLIPCAIDODPYFRMTDVAAPRIGYKPPALHSTFFPALQAG 301
DB 293 FSNSEFPQIFRRTDIOCLIPCAIDODPYFRMTDVAAPRIGYKPPALHSTFFPALQAG 352
QY 302 KMSASDPNSSIFLDTAKQIKTKVKNKHAFFSGGRDTEEHROFGNCDDVDSFMYLTFE 361
DB 353 KMSASDPNSSIFLDTAKQIKTKVKNKHAFFSGGRDTEEHROFGNCDDVDSFMYLTFE 412
QY 362 DDDKLEQIRKDYTGAMLTGELKALIEVLOPLIAEHQARKEVDTEYKEFMPRKLS 421
DB 413 DDDKLEQIRKDYTGAMLTGELKALIEVLOPLIAEHQARKEVDTEYKEFMPRKLS 472
QY 422 DFO 424
DB 473 DFO 475

```

## RESULT 3

tryptophan--tRNA ligase (EC 6.1.1.2) [validated] - rabbit

N:Alternate names: tryptophanyl-tRNA synthetase

C:Species: Oryctolagus cuniculus (domestic rabbit)

C:Date: 30-Sep-1992 #sequence\_revision 13-Feb-1998 #text\_change 26-May-2000

C:Accession: A35904; S37396

R:Lee, C.C.; Craigen, W.J.; Muzny, D.M.; Harlow, E.; Caskey, C.T.

Proc. Natl. Acad. Sci. U.S.A. 87, 3508-3512, 1990

A:Title: Cloning and expression of a mammalian peptide chain release factor with sequ

A:Reference number: A35904; MUID:90239043

A:Accession: A35904

A:Molecule type: mRNA

A:Residues: 1-475 <LEF>

A:Cross-references: GB:M33460

A:Experimental source: Dalphin, M.E.; Justesen, J.; Powell, R.J.; Drugeon, G.; McCaughan, K

EMBO J. 12, 4013-4019, 1993

A:Title: Mammalian polypeptide chain release factor and tryptophanyl-tRNA synthetase

A:Reference number: S37396; MUID:94009008

A:Accession: S37396

A:Molecule type: mRNA

A:Residues: 166-177 <FRO>

C:Genetics:

A:Gene: WRS

C:Complex: homodimer [validated, MUID:94009008]

A:Description: EC 6.1.1.2 [validated, MUID:94009008]; catalyzes the ATP-dependent for

A>Note: mammalian WRS (tryptophanyl-tRNA synthetase) and eRF (polypeptide chain rele

C:Superfamily: mammalian tryptophan--tRNA ligase; amino acid--tRNA ligase repeat homo

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: October 24, 2002, 12:49:45 ; Search time 14.2521 Seconds  
(without alignments)  
2858.658 Million cell updates/sec

Title: US-09-813-718-10\_COPY\_48\_471

Perfect score: 2246  
Sequence: 1 MSYKAAAGEYKADCPGPNP.....VTDEIVKEFMTPRKLSFDQ 424

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 28138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 28138

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

## Database :

PIR\_71:.\*  
1: p1r1:.\*  
2: p1r2:.\*  
3: p1r3:.\*  
4: p1r4:.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2231	99.3	471	1 A41706	tryptophan--trna 1
2	2134.5	95.0	475	1 YWBO	tryptophan--trna 1
3	2041	90.9	475	1 TWBPR	tryptophan--trna 1
4	2032	90.5	481	2 S50053	tryptophan--trna 1
5	1210	53.9	395	2 S58157	hypothetical prote
6	1163	51.8	432	2 S51901	tryptophan--trna 1
7	907	40.4	385	2 C90190	tryptophan--trna 1
8	803	35.8	385	2 C75020	tryptophan--trna 1
9	626.5	27.9	301	2 G71206	tryptophan--trna 1
10	534.5	23.8	380	2 G84373	tryptophan--trna 1
11	409.5	18.2	370	2 F64476	tryptophan--trna 1
12	397.5	17.7	364	2 E69131	tryptophan--trna 1
13	386	17.2	134	2 T43806	tryptophan--trna 1
14	370.5	16.5	420	2 E69451	tryptophan--trna 1
15	353	15.7	374	2 D72477	probable tryptoph
16	274.5	12.2	513	2 F84371	tryptophan--trna 1
17	192	8.5	323	2 H69346	tyrosyl--trna synth
18	188	8.4	364	2 E72512	probable tyrosyl-t
19	185.5	8.3	341	2 D95260	tryptophan--trna 1
20	185.5	8.3	341	2 G98125	tryptophan--trna 1
21	174.5	7.8	341	2 B86633	tryptophan--trna 1
22	172	7.7	394	2 A45999	tyrosine--trna 1
23	170.5	7.6	366	2 A45999	tyrosine--trna 1
24	169	7.5	395	2 H70385	tyrosine--trna 1
25	167	7.4	351	2 E75438	tryptophan--trna 1
26	166.5	7.4	346	2 B71496	tryptophan--trna 1
27	164.5	7.3	895	2 A86410	protein F3M18.22
28	163.5	7.3	353	2 E70100	tryptophan--trna 1
29	163	7.3	460	2 C84750	probable tyrosyl-t

30	159.5	7.1	337	2 F71300	tryptophan--trna 1
31	158.5	7.1	375	2 B75072	tyrosyl--trna synth
32	158	7.0	346	2 C81654	tryptophan--trna 1
33	154.5	6.9	102	2 T74994	probable tryptoph
34	152	6.8	344	2 H86590	tryptophan--trna 1
35	152	6.8	344	2 C72034	tryptophan--trna 1
36	149.5	6.7	327	2 C84374	tyrosyl--trna synth
37	149.5	6.7	408	2 T03741	probable tyrosine-
38	146	6.5	339	2 E64676	tryptophan--trna 1
39	146	6.5	375	2 F71093	tyrosine--trna 1
40	145	6.5	337	2 A11066	tryptophan--trna 1
41	138	6.1	319	2 H69102	tyrosine--trna 1
42	135.5	6.0	328	2 C72370	tryptophan--trna 1
43	135	6.0	365	2 E82052	tryptophan--trna 1
44	132.5	5.9	343	2 S73024	tryptophan--trna 1
45	132	5.9	328	1 YWBSF	tryptophan--trna 1

## ALIGNMENTS

## RESULT 1

A41706  
tryptophan--trna 1 ligase (EC 6.1.1.2) [similarity] - human  
N:Alternate names: Interferon-Inducible protein IFP53; peptide-chain release factor h  
C:Species: Homo sapiens (man)  
C:Date: 19-May-2000 #sequence, revision 19-May-2000 #text, change 19-May-2000  
C:Accession: A41633; A41706; S19246; JN0676; JH0533; S26287  
R: Fleckner, J.; Rasmussen, H.H.; Justesen, J.  
Proc. Natl. Acad. Sci. U.S.A. 88, 11520-11524, 1991  
A:Title: Human interferon gamma potentially induces the synthesis of a 55-kDa protein (9  
A:Reference number: A41633; MUID:92107982  
A:Accession: A41633  
A:Status: Preliminary  
A:Molecule type: mRNA  
A:Residues: 1-471 <RUB>  
A:Cross-references: GB:X59892; NID:930820; PIDN:CAA42545.1; PID:930821  
R: Rubin, B.Y.; Anderson, S.L.; Xing, L.; Powell, R.J.; Tate, W.P.  
J. Biol. Chem. 266, 24245-24248, 1991  
A:Title: Interferon induces tryptophanyl-trna synthetase expression in human fibrobla  
A:Reference number: A41706; MUID:92105071  
A:Accession: A41706  
A:Status: Preliminary  
A:Molecule type: mRNA  
A:Residues: 1-471 <RUB>  
A:Cross-references: GB:M77804; NID:9184656; PIDN:AAA67324.1; PID:9184657  
R: Buwitt, U.; Fliohr, T.; Boeltger, E.C.  
EMBO J. 11, 489-496, 1992  
A:Title: Molecular cloning and characterization of an interferon induced human cDNA w  
A:Reference number: S19246; MUID:92164636  
A:Accession: S19246  
A:Status: Preliminary  
A:Molecule type: mRNA  
A:Residues: 1-423, 'R', 425-471 <BUW>  
A:Cross-references: EMBL:X62570; NID:932708; PIDN:CAA44450.1; PID:932709  
A:Note: 213-Ser and 214-Tyr were also found  
R: Frolova, L.Y.; Grigorjeva, A.Y.; Sudomolina, M.A.; Kisselev, L.L.  
Gene 120, 237-245, 1993  
A:Title: The human gene encoding tryptophanyl-trna synthetase: Interferon-response el  
A:Reference number: JN0676; MUID:93292992  
A:Accession: JN0676  
A:Molecule type: DNA  
A:Residues: 1-141, 182-471 <PRO1>  
A:Cross-references: GB:X67918; GB:S62837; NID:937968; GB:X67919; NID:937969; GB:X6792  
4; NID:937974; GB:X67925; GB:S62855; NID:937975; GB:X67926; GB:S62856; NID:937976; GB  
A:Note: the authors translated the codon GGG for residue 55 as Cys and GAG for residu  
R: Frolova, L.Y.; Sudomolina, M.A.; Grigorjeva, A.Y.; Zhirovleva, O.L.; Kisselev, L.L.  
Gene 109, 291-296, 1991  
A:Title: Cloning and nucleotide sequence of the structural gene encoding for human tr  
A:Reference number: JH0533; MUID:92112058  
A:Accession: JH0533  
A:Molecule type: mRNA





Fri Oct 25 09:23:05 2002

us-09-813-718-10\_copy\_48\_471.ra

Page 9

Search completed: October 24, 2002, 12:54:33  
Job time : 11.3652 secs

---



TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-928-100-4

Query Match 4.8%; Score 108; DB 3; Length 197;  
Best Local Similarity 25.3%; Pred. No. 0.00081;  
Matches 42; Conservative 27; Mismatches 73; Indels 24; Gaps 6;

QY 260 IPCAIDDDPYFRMTRD-----VAPRIGYKPKALLHSFFPALOGAQTMSA 305  
DB 5 VPVGTDDKPMIEQIREIVRSFNNAVNCVDLYVEPGIYPENE--RAGRLPGIDG--NAKMSK 61  
QY 306 SDPNSSIFLDTAKOIKTKVKNKHAFFSGGRDTIEHRQFGNCVDVSMYLTFF--LEDD 363  
DB 62 S-LNNGIYLDADDTLKKKWSMTDPDHI RVEDPGKIEGN---WVFHYLDVGRPEDA 116  
QY 364 DKLEQIRKDYTSGAMLTGELKALIEVLOPLIAHQARRREYVDEI 409  
DB 117 QEIADMERYQRGGLGVKTKRYLLEIRERLGPRIERRIEFAMK 162

## RESULT 12

US-09-183-134-2  
Sequence 2, Application US/09183134  
Patent No. 6165759  
GENERAL INFORMATION:  
APPLICANT: Lawlor, Elizabeth  
TITLE OF INVENTION: NOVEL TRYPTOPHANYL TRNA SYNTHETASE  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Smithkline Beecham Corporation  
STREET: 709 Swedeland Road  
CITY: King of Prussia  
STATE: PA  
COUNTRY: USA  
ZIP: 19406-0939  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/183,134  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/923,867  
FILING DATE: 04-SEP-1997  
APPLICATION NUMBER: 9619072.3  
FILING DATE: 12-SEP-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Gimmli, Edward R.  
REGISTRATION NUMBER: 38,891  
REFERENCE/DOCKET NUMBER: P31624  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-270-4478  
TELEFAX: 610-270-5090  
TELEX:  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 197 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-183-134-2

Query Match 4.8%; Score 108; DB 4; Length 197;  
Best Local Similarity 25.3%; Pred. No. 0.00081;  
Matches 42; Conservative 27; Mismatches 73; Indels 24; Gaps 6;  
QY 260 IPCAIDDDPYFRMTRD-----VAPRIGYKPKALLHSFFPALOGAQTMSA 305  
DB 5 VPVGTDDKPMIEQIREIVRSFNNAVNCVDLYVEPGIYPENE--RAGRLPGIDG--NAKMSK 61

DB 5 VPVGTDDKPMIEQIREIVRSFNNAVNCVDLYVEPGIYPENE--RAGRLPGIDG--NAKMSK 61  
QY 306 SDPNSSIFLDTAKOIKTKVKNKHAFFSGGRDTIEHRQFGNCVDVSMYLTFF--LEDD 363  
DB 62 S-LNNGIYLDADDTLKKKWSMTDPDHI RVEDPGKIEGN---WVFHYLDVGRPEDA 116  
QY 364 DKLEQIRKDYTSGAMLTGELKALIEVLOPLIAHQARRREYVDEI 409  
DB 117 QEIADMERYQRGGLGVKTKRYLLEIRERLGPRIERRIEFAMK 162

## RESULT 13

US-09-492-581-4  
Sequence 4, Application US/09492581  
Patent No. 6346409  
GENERAL INFORMATION:  
APPLICANT: Gentry, Danlle  
APPLICANT: Greenwood, Claire  
APPLICANT: Lawlor, Elizabeth  
TITLE OF INVENTION: No. 6346409el trps  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Smithkline Beecham Corporation  
STREET: 709 Swedeland Road  
CITY: King of Prussia  
STATE: PA  
COUNTRY: USA  
ZIP: 19406-0939  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/492,581  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/928,100  
FILING DATE: 12-SEP-1997  
APPLICATION NUMBER: 9619072.3  
FILING DATE: 12-SEP-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Gimmli, Edward R.  
REGISTRATION NUMBER: 38,891  
REFERENCE/DOCKET NUMBER: P31624-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-270-4478  
TELEFAX: 610-270-5090  
TELEX:  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 197 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-492-581-4

Query Match 4.8%; Score 108; DB 4; Length 197;  
Best Local Similarity 25.3%; Pred. No. 0.00081;  
Matches 42; Conservative 27; Mismatches 73; Indels 24; Gaps 6;  
QY 260 IPCAIDDDPYFRMTRD-----VAPRIGYKPKALLHSFFPALOGAQTMSA 305  
DB 5 VPVGTDDKPMIEQIREIVRSFNNAVNCVDLYVEPGIYPENE--RAGRLPGIDG--NAKMSK 61  
QY 306 SDPNSSIFLDTAKOIKTKVKNKHAFFSGGRDTIEHRQFGNCVDVSMYLTFF--LEDD 363  
DB 62 S-LNNGIYLDADDTLKKKWSMTDPDHI RVEDPGKIEGN---WVFHYLDVGRPEDA 116  
QY 364 DKLEQIRKDYTSGAMLTGELKALIEVLOPLIAHQARRREYVDEI 409  
DB 117 QEIADMERYQRGGLGVKTKRYLLEIRERLGPRIERRIEFAMK 162



PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/705,868  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0117 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 388 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
IMMEDIATE SOURCE:  
LIBRARY: GenBank  
CLONE: 1184699  
US-09-123-615-4

Query Match 5.98; Score 132.5; DB 3; Length 388;  
Best Local Similarity 21.2%; Pred. No. 4,7e-06;  
Matches 79; Conservative 64; Mismatches 149; Indels 80; Gaps 19;

QY 83 HFLRGCI--FESHDMNVLDAYEKKRFVLYTGSGPSEAMHGHILPFI--FTKWLQDV 139  
DB 12 HLIRNLOEVLGEERKILKERELK---IYWGATATGKP-HAAYVPMKIDFLKAG 66  
QY 140 FNVPLVI---QMTDDEKYLKDLTDQAYGDAVENAKDIACGFDINKTFISDLDMG 195  
DB 67 CEVITLFDLHAYLDNMKAPWELLETLRYSYENVIKAM-LESIGVPLEKLFKIGTDQ- 124  
QY 196 MSSGFYKAVKVIQKHVFNQVYK---GTFGTTSDSCIGKISFPALQAPSPNSPQIFR 251  
DB 125 LSKETTLIDVYRLSSVYQHDSSKKAQAEVVKQVEHPLLSGLLYPGIQA---LDEEYLVK- 179  
QY 252 DRTDIOCLPCALDODPFRMTRDVAPRIGYKPKALLHSTFPALOGQTKMSASDPSS 311  
DB 180 ---DAQF---GGIDQRKLTFAEKLTPALGTSKRVHLMNPMVPLTG--SKMSSSEESK 231  
QY 312 IFLDTAKOIKTKYKNKNAF-----SGGRDTIEEH-----ROFGNCDVD 350  
DB 232 IDLDRKEDVAKLKK-AFCERGVNENNGVSLFKHVLFPKSEFVILRDKMGDN--- 286  
QY 351 VSEMYLFFLEDDDKLEQIRKDYTSGLAMLTGELKALIEVLOPLAEHQARKKEVTDEIV 410  
DB 287 ---RTYAYVD-----LEKDFAAEVVHPGDLKNSVEVALNKLD-----DPFR 325  
QY 411 KEFMP--RKLS 420  
DB 326 KEFMPALKLA 337

RESULT 8  
US-08-855-910-11  
Sequence 11, Application US/08855910  
Patent No. 6221640  
GENERAL INFORMATION:  
APPLICANT: Tao, Jianshi  
APPLICANT: Sassanfar, Mandana  
APPLICANT: Gallant, Paul L.  
APPLICANT: Shen, Xiaoyu  
APPLICANT: Avriuch, Anthony S.  
APPLICANT: Yu, Russell V.  
TITLE OF INVENTION: ENTEROCOCCAL AMINOACYL-tRNA SYNTHETASE  
TITLE OF INVENTION: PROTEINS, NUCLEIC ACIDS AND STRAINS COMPRISING SAME  
NUMBER OF SEQUENCES: 71  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
STREET: Two Militia Drive

CITY: Lexington  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02173  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/855,910  
FILING DATE: 14-MAY-1997  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Brook, David E.  
REGISTRATION NUMBER: 22,592  
REFERENCE/DOCKET NUMBER: CP195-08  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (781) 861-6240  
TELEFAX: (781) 861-9540  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 418 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-855-910-11

Query Match 5.18; Score 113.5; DB 4; Length 418;  
Best Local Similarity 20.5%; Pred. No. 0.00069;  
Matches 82; Conservative 47; Mismatches 112; Indels 159; Gaps 18;

QY 111 LYTGRGSEAMHGHILPFI--FTKWLQDVENVPLVI----- 146  
DB 33 LYCGVDPGSDMHGHILPFI--FTKWLQDVENVPLVI----- 92  
QY 147 -----QMTDDEKYLW-KDLTLDQAYGDAVEN-----AKDIIA 177  
DB 93 AVQHNVDLSNQMKKLPKQDAEVTMVNNDLSLDFLRDVGKNNVMTLAKDIVA 152  
QY 178 CGFDINKTFISDLDMGMS--GFYKAVKIQKHVFNQVYK----- 218  
DB 153 S--RLESGISFTFTYQIQLSIDFY---TLHKKNIIQLQIGADQWGNITAGLDIRKE 207  
QY 219 -----JEGFT-----DSDCIGKISFPALQAPSPNSP--QIFRRTDIOCLPCAI 264  
DB 208 GPEAKVGLTLPMLKADGTGFKTAGALWLDPRKTSFFEFYQFWLNQDD----- 258  
QY 265 DQDVFRTMRDVARPRIGYKPKALLHSTFPALOGAQTMSASDPNSSIFLTDTAKOIKTK 324  
DB 259 -----RDV---IKYLK-----FFFLDKEE-----IDALAEKVEKE 286  
QY 325 VNKAHFGGRDTIEEHROFGNCDVDVSEMYLFFLEDDDKLEQIRKDYTSGLAMLTGELK 384  
DB 287 PKR--EAQRRLAE-----VTRFVHDALAEBAOK--ISEALFSGNIK 326  
QY 385 KALIEVLQ-----PLAEHQARKKEVTDEIVKEFMPRK 418  
DB 327 DLTIEELQGLEHVPTV-ETTKAKNIVDWLVTIEIEPSK 365

RESULT 9  
US-09-352-990-28  
Sequence 28, Application US/09352990  
Patent No. 6255090  
GENERAL INFORMATION:  
APPLICANT: Famodu, Layo O.  
APPLICANT: Orozco, Buddy  
APPLICANT: Rafalski, Antoni  
TITLE OF INVENTION: Plant Aminoacyl-tRNA Synthetase  
FILE REFERENCE: BB-1191  
CURRENT APPLICATION NUMBER: US/09/352,990  
CURRENT FILING DATE: 1999-07-14

INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 409 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-743-130A-2

Query Match 7.2%; Score 162.5; DB 2; Length 409;  
Best Local Similarity 20.7%; Pred. No. 2.3e-09;  
Matches 78; Conservative 59; Mismatches 132; Indels 107; Gaps 16;

OY VLDAYENK-KRPYLYTGRGSSSEAMHVGHLIPFI-----FTKWLQDV-----F 140  
DB 27 IKDYLEKEKRPKIYWGTAFTGK-P-HCGYFVPMIKLAHLKACCEVYTLADLHAFIDNM 85  
OY 141 NVPLVIOMTDEKYLWKDLTLDQAYGDAVENAKDIACGFDINKTFIFSDLDYMGSSGF 200  
DB 86 KAPLEVYKRAKYVEFVYKAILKINSVPIERLKFVYSSYQKGDYV--MDLFRKLSNIV 142  
OY 201 YKNVYK-----IQKVTENQVKGIFGFTDSCIGKISFPAIOAPFSNSFPQIRDRID 255  
DB 143 SONDAKRRAGADYVKQVNPDLLESLI-----YPLMQA-----IDEBHLG 180  
OY 256 IOCLIPCAIDODPYFRMTDVAPRIGYKPKPALHSTFPALOGAQTAKMSASDPNSSIFLT 315  
DB 181 VDAQSG-GVDOKKITVLAENLPSIGYKRAHLNMPVPGI-GQGGKMSASDPNSKIDIT 238  
OY 316 DTAKOIKTRVKNHAFSGG--RDT-----IEEHROFGN 346  
DB 239 EEPKVVKKKNSAYCAPGLKNDGLIAFIEVYIOPIAELKTGVEGAFKIDLPREKYG- 297  
OY 347 CDVYSEPMITLTFLEDDDKLEQIRKDYSGAMLTGELKALL-----EYIOLPIAEHQARR 402  
DB 238 ---PLSY-----SIEQLKADFDVGKLPAPDLKLGADKINELLAPIRAEFESS- 343  
OY 403 KEVTDEIVEFMTPRK 418  
DB 344 -----EEFYVAOK 351

## RESULT 6

US-08-705-868-4

Sequence 4, Application US/08705868

Patent No. 5885798

GENERAL INFORMATION:

APPLICANT: Bandman, Olga

APPLICANT: Coleman, Roger

APPLICANT: Au-Young, Janice

APPLICANT: Murry, Lynn E.  
TITLE OF INVENTION: NOVEL MONOCYTE ACTIVATING CYTOKINE

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSEE: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive

CITY: Palo Alto

STATE: CA

COUNTRY: U.S.

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ Version 1.5

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/705,868

FILING DATE: Filed Herewith

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0117 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:

LENGTH: 388 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
IMMEDIATE SOURCE:  
LIBRARY: Genbank  
CLONE: 1184699  
US-08-705-868-4

Query Match 5.9%; Score 132.5; DB 2; Length 388;  
Best Local Similarity 21.2%; Pred. No. 4.7e-06;  
Matches 79; Conservative 64; Mismatches 149; Indels 80; Gaps 19;

OY 83 HFLRRGI--FFSHRDMNOVLDAVENKKRPYLYTGRGSSSEAMHVGHLIPFI-FTKWLQDV 139  
DB 12 HLTIRNLQEVLEGEKELKELKEREIK-----IYWGTAITGK-P-HVAYFVPMKIAIDFLKAG 66  
OY 140 FNVPLVI-----QMTDEKYLWKDLTLDQAYGDAVENAKDIACGFDINKTFIFSDLDYMG 195  
DB 67 CEVTILFADLHAYLDNMKAPWELLELRVSYYENVIKAM-LESIGVLEKLEKFGCTDYO- 124  
OY 196 MSSGYKKNVVKIQKVTENQV-----GIFGFTDSCIGKISFPAIOAPFSNSFPQIR 251  
DB 125 LSKEYTLVDYRLSSVYQTHDSKKAQAEVVKQVHEHPLLSGLLYPGIDA--LDEEYLVK-- 179  
OY 252 DRTDIOCLIPCAIDODPYFRMTDVAPRIGYKPKPALHSTFPALOGAQTAKMSASDPNSS 311  
DB 180 ---DAQF---GIDQKRIKTFEAKYLPALGYSKRHLNMPVPGITG--SKMSSEESK 231  
OY 312 IFLVTPAKOIKTRVKNHAF-----SGGRDIEEH-----ROFGNCVD 350  
DB 232 IDLLRKEVDKKKLKK-AFCEPGVNEVNGVLSFIKHVLPFLKSEFVILRDEKMGN--- 286  
OY 351 VSFMYLTFLEDDDKLEQIRKDYSGAMLTGELKALLIEVYIOLPIAEHQARKEVTDIV 410  
DB 287 ---KTYTAYVD-----LEKDAEAVVHPGDLKNSVEVALKLL-----DPIR 325  
OY 411 KEFMTPR--RKLS 420  
DB 326 EKFNTPALKKILA 337

## RESULT 7

US-09-123-615-4

Sequence 4, Application US/09123615

Patent No. 6090377

GENERAL INFORMATION:

APPLICANT: Bandman, Olga

APPLICANT: Coleman, Roger

APPLICANT: Au-Young, Janice

APPLICANT: Murry, Lynn E.  
TITLE OF INVENTION: NOVEL MONOCYTE ACTIVATING CYTOKINE

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSEE: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive

CITY: Palo Alto

STATE: CA

COUNTRY: U.S.

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ Version 1.5

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/123,615

FILING DATE:

US-09-492-581-2

Query Match	8.3%;	Score 185.5;	DB 4;	Length 341;
Best Local Similarity	24.0%;	Pred. No. 4.7e-12;		
Matches	81;	Conservative	60;	Mismatches 128;
			Indels	69;
			Gaps	15;

```

QY 107 KPEFLYGRGSSAMVGHILPIETFKYLODVENVLYLOMDDEKY-LMKLOUTLOAY 165
Dp 3 KPIIL-TGDRYTCG-LHIGHVSLKRR-----VLOEEDKYDMFVYLAQOAL 49

QY 166 GDAVEN-----AKDIIACGPIINTFFI--SDLDYMGSSGFYKNV---KIOK 209
Dp 50 TDAKKDDPOTIVESIGNVALLYLAAGDIPNKSJTIFIOQPELWLSMYVNYVLSTARLE 109

QY 210 HATFNQKGIENGFTDSDCIGKISFPALQAA--PFSNSFPQIFRDRDIOCLIPCAIDOP 267
Dp 110 NPVTKEISQGFESIPTGFLVYRINQADITAFKANY-----VPVGIDK 156

QY 268 PYFKMTD-----VAPRIQYKRPALHSTFFPALQOATKKSASADPNSSITF 313
Dp 157 PMIEOTREIYNSFNNAAYNCVULVEPEGIYENE--RAGRLPGIDG-NAKSKS-LNNGITV 212

QY 314 LTFDAKOIKTVKNKHAFGSGHDTIEEHRQGGNCDVVSFMYLTF--LEDDDKLEIQIK 371
Dp 213 LADDAOTLRKKVSMYTDPRHIVEDPQKIEGN---MVFNYLUDFQGRPDAGIEMDK 268

QY 372 DYTSGAMITGELKALLLEVLOPLAEHQARKKEVTDEI 409
Dp 269 RYQRGSGADVTKRYLLETERELRIGPIRRERIEFAKDM 306

```

#### RESULT 4

US-08-743-130A-39  
; Sequence 39, Application US/08743130A

```

GENERAL INFORMATION:
APPLICANT: Sassanfar, Mandana
APPLICANT: Gallant, Paul L.
APPLICANT: Shen, Xiaoyu
APPLICANT: Tao, Nianjun
APPLICANT: Tao, Jianshi
APPLICANT: Houman, Farida
TITLE OF INVENTION: CANDIDA TYROSYL-TRNA SYNTHETASE
TITLE OF INVENTION: PROTEINS, NUCLEIC ACIDS AND STRAINS COMPRISING SAME
NUMBER OF SEQUENCES: 41
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P. C.
STREET: Two Milltia Drive
City: Lexington
STATE: Massachusetts
COUNTRY: USA
ZIP: 02173
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/743,130A
FILING DATE: 01-NOV-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Brook Esq., David E.
REGISTRATION NUMBER: 22,592
TELECOMMUNICATION INFORMATION:
REFERENCE/DOCKET NUMBER: CP195-12
TELEPHONE: (617) 861-6240
TELEFAX: (617) 861-9540
INFORMATION FOR SEQ. ID NO.: 39:
SEQUENCE CHARACTERISTICS:
LENGTH: 409 amino acids
TYPE: amino acid
STRANDEDNESS:

```

US-08-743-130A-39

Query Match	7.38	Score 164.5	DB 2	Length 409
Best Local Similarly	20.78	Pred. No. 1.4e-09		
Matches 78	Conservative 59	Mismatches 132	Indels 107	Gaps 16

```

QY 99 VLDVAENK-KRFVLYTGTGSPSEAMHGHILPFI-----ETKMLDQV-----F 140Q
Db 27 IKDVLKERNRKYKLYMGTAIPGKR-HCGYFPMPIKTLAHLKAGEYVYLLADLAFLADNM 85
QY 141 NVPLVIOATDDEKYLKMDLLTLDQAYGAVENAKDIIACGFDFINKFTFSDDLGYKMSGGF 200Q
Db 86 KALEVYVYRAKYEEFVYKALIKLSINWPIERLTKFVYGSSYQGGDQV---MDLFKLSINIV 142Q
QY 201 YKNVYK-----IQKHVTFPNQYKGIJFGFLDSDCIKISFPALQAAPESSNSFPQJFRORTD 255Q
Db 143 SQNDAKRAGADVYQVAPNPLLSGLI-----YPLMQA-----IDEBHLG 180Q
QY 256 IQCLIPCAIDDDPFYRMTRDVAPRIGPKPALLHSTFEPFALOGAQTQKMSASDPNSJFELT 315Q
Db 181 VDAQFG-GVDQRKTFVLAEEMLPISITGKRAHLMNPMVPGI-GGGKMSASDPNSKIDIT 238Q
QY 316 DTAAQIKTKYVKNHAFSGS-RDT-----IEEHROFGCN 346Q
Db 239 EEPKYVKKKVAASACARGELKNDNLIAFIEVQIAPIELKTGVGAFKRLDIDREKXIGG- 297Q
QY 347 CDVDVSEMYTLTFLEDDDKLEQIKRDYTSGAMLTGELKKALI---EVLQPLAIEHQARR 402Q
Db 298 ---PLSY-----DSIQFLKADFEVCGKLAPPLDLSGVADKINETLAPIRAFEPS- 343Q
QY 403 KEVYDEIYKEFMFPRK 418
Db 344 -----EEFYVAOK 351

```

## RESULT 5

US-08-743-130A-2  
; Sequence 2, Application US/08743130A

GENERAL INFORMATION:  
APPLICANT: Sassanfar, Mandana  
APPLICANT: Gallant, Paul L.  
APPLICANT: Shen, Xiaoyu  
APPLICANT: Tao, Nianjun  
APPLICANT: Tao, Jianshi  
APPLICANT: Houman, Fariba  
TITLE OF INVENTION: CANDIDA TYROSYL-tRNA SYNTHETASE  
TITLE OF INVENTION: PROTEINS, NUCLEIC ACIDS AND STRAINS COMPRISING SAME  
NUMBER OF SEQUENCES: 41  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
STREET: Two Millitia Drive  
CITY: Lexington  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02173  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/743,130A  
FILING DATE: 01-NOV-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Brook Esq., David E.  
REGISTRATION NUMBER: 22,592  
REFERENCE/DOCKET NUMBER: CP195-12  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 861-6240  
TELEFAX: (617) 861-9540



Qy	87	RGIFEHSDHMQVDAVENKPKFELTYGRGSSSAMVGHIPITFKMLQDFNVPLVI	146
Db	73	RGVEFSQMDLRILDIKVENHGFPEFLTYGRGSSDSMLGHMVPRITFKMLQDFNVPLVI	132
Qy	147	QMTDEKYLAK-DLTLDQAVSGDAVENAKDIACGFCFDINKTFIISDLDGMGSSGFKANVY	205
Db	133	ELTDEKELFYHQTLLIDYKCFEAEKKNDIIAVGFENEFIFSDLDYMG--GAFYENVV	190
Qy	206	KIQKHVFENQYKGFJGFSDSCIKISFPATQAPSFNSFPQTRPRTQICLPCAI	265
Db	191	RTSQRTITSTAKAVEGFFDSDCKIHASIQATAPRSPDPVGLGPRTPCLPCAI	250
Qy	266	ODPFKRMTRDVAPIRGYKPKYALLSTFFPALQSGQTCKMSADPNSSIFLDTAKQIKTV	325
Db	251	QDPFRRCRDVADLRFLPKPALIHAKFFPALQGSTKMSADPTTSTFMSDTRAKQIKTI	310
Qy	326	NKHAFFSGGRDYTEEHROFGNCDVDFVMYTFLEDDDKLEQIRKDYTGAMLTGLK	385
Db	311	NKYAFSGGRATAEEHRELGTGNEVDVAFOYLSPFSYDDEKLAQLQCYGRGEIISGEMK	370
Qy	386	ALIEVLOPLIAHQARKREYDELVEKMPRKLSF	421
Db	371	ECITVLOEFAVSAYOERKSDQVVEKFMKPKLVE	406

```

RESULT 2
US-08-928-100-2
: Sequence 2, Application US/08928100
: Patent No. 6046174
: GENERAL INFORMATION:
: APPLICANT: Gentry, Danielle
: APPLICANT: Greenwood, Claire
: APPLICANT: Lawlor, Elizabeth
: TITLE OF INVENTION: No. 6046174el trps
: NUMBER OF SEQUENCES: 6
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: SmithKline Beecham Corporation
: STREET: 709 Swedeland Road
: CITY: King of Prussia
: STATE: PA
: COUNTRY: USA
: ZIP: 19406-0939
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FASTSEQ for Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/928,100
: FILING DATE: 12-SEP-1997
: CLASSIFICATION: 424
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 9619072.3
: FILING DATE: 12-SEP-1996
: ATTORNEY/AGENT INFORMATION:
: NAME: Gimmil, Edward R
: REGISTRATION NUMBER: 38,891
: REFERENCE/DOCKET NUMBER: P31624-1
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 610-270-4478
: TELEFAX: 610-270-5090
: TELEX:
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 341 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: protein
:
: US-08-928-100-2
Query Match      8.3%; Score 185.5; DB 3; Length 341.
Best Local Similarity 24.0%; Prid. No. 4,7e-12;

```

	Matches	81,	Conservative	60;	Mismatches	128;	Indels	69;	Gaps	15;
QY	107	KPEFYLYTGRGSSSEAMHVGHLPIFTKWLQDENVNPLVIQTMDDEKY-LMKDLTLDOAY	165							
Db	3	KPIL-TGDPRITGK-LTHGHVYGSLKNR-----VLLOEEDSKYDMFEFLDAQAL	49							
QY	166	GDAVEN-----AKDIACGFDDINKPTIF--SDLDPMGNSSGGYKNV---KIOK	209							
Db	50	TDHADDPOTIVESIGNVALDYLVAGLNDKSTIFIQSIOIPELAELSMYYMNLVSIAKLAR	109							
QY	210	HVTENQNVGIRFGFTDSDIGIKISFPAIOAA--PSFSNSFPQIFRDRTDIOCLIPCAIDOD	267							
Db	110	NPTVTELSQNGFGSGSIPTGLVLYIAQAADITARKAY-----VPVGTQOK	156							
QY	268	PYFKMRTRD-----VAPRICPKPALHSFFPALOGAQTKMSASDPNSSIF	313							
Db	157	PMIEOTREIVSFNNNAVCODYLVEPEGITYPENR--RAQRLLPDGS-NARKMSKS-LNMGCI	212							
QY	314	LTPDTKOIKRTVYNKHAFSGGRDITEHHQPGNCNOVDVSPFWLTFF--LEDDDKLEQIRK	371							
Db	213	LADDADTLRKRYKMSMYTDPDHLRVDPCKISEN---WVFHLDVFGRPEDQAQETADMK	268							
QY	372	DYTSGAMLTELKALKALIEVLOPLINEHQARREVTVDEI	409							
Db	269	KYRGGLDGVDVKTKRYLLEIRELEGPIERKRILEFAKDM	306							

```

1      RESULT 3
2      US-09-492-581-2
3      :
4      : Sequence 2, Application US/09492581
5      : Patent No. 6346409
6      :
7      : GENERAL INFORMATION:
8      :
9      : APPLICANT: Gentry, Danile
10     : APPLICANT: Greenwood, Claire
11     : APPLICANT: Lawlor, Elizabeth
12     : TITLE OF INVENTION: No. 6346409el tips
13     : NUMBER OF SEQUENCES: 6
14     :
15     : CORRESPONDENCE ADDRESS:
16     : ADDRESSEE: SmithKline Beecham Corporation
17     : STREET: 709 Swedeland Road
18     : CITY: King of Prussia
19     : STATE: PA
20     :
21     : COUNTRY: USA
22     : ZIP: 19406-0939
23     :
24     : COMPUTER READABLE FORM:
25     : MEDIUM TYPE: Diskette
26     : COMPUTER: IBM Compatible
27     : OPERATING SYSTEM: DOS
28     : SOFTWARE: FASTSEQ for Windows Version 2.0
29     :
30     : CURRENT APPLICATION DATA:
31     : APPLICATION NUMBER: US/09/492,581
32     : FILING DATE:
33     :
34     : CLASSIFICATION:
35     :
36     : PRIOR APPLICATION DATA:
37     : APPLICATION NUMBER: US/08/928,100
38     : FILING DATE: 12-SEP-1997
39     : APPLICATION NUMBER: 96119072.3
40     : FILING DATE: 12-SEP-1996
41     :
42     : ATTORNEY/AGENT INFORMATION:
43     : NAME: Gimmli, Edward R
44     : REGISTRATION NUMBER: 38,991
45     : REFERENCE/DOCKET NUMBER: P31624-1
46     : TELECOMMUNICATION INFORMATION:
47     : TELEPHONE: 610-270-4478
48     : TELEFAX: 610-270-5090
49     :
50     : TELEX:
51     :
52     : INFORMATION FOR SEQ ID NO: 2:
53     :
54     : SEQUENCE CHARACTERISTICS:
55     : LENGTH: 341 amino acids
56     : TYPE: amino acid
57     : STRANDEDNESS: single
58     : TOPOLOGY: linear
59     :
60     : MOLECULE TYPE: protein
61     :
62     :

```

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: October 24, 2002, 12:49:45 ; Search time 10.3652 Seconds  
(without alignments)  
999.159 Million cell updates/sec

Title: US-09-813-718-10\_COPY\_48\_471

Perfect score: 2246  
Sequence: 1 MSYKAAAGEDYKADCPGPNP.....VTDEIVKEFWTPRKLSDFQ 424

Scoring table: BLOSUM62  
Gap 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/PCUS.COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1218.5	54.3	424	US-08-876-885-26	Sequence 26, Appl
2	185.5	8.3	341	US-08-928-100-2	Sequence 2, Appl
3	185.5	8.3	341	US-09-492-581-2	Sequence 2, Appl
4	164.5	7.3	409	US-08-743-130A-39	Sequence 39, Appl
5	162.5	7.2	409	US-08-743-130A-2	Sequence 2, Appl
6	132.5	5.9	388	US-08-705-868-4	Sequence 4, Appl
7	132.5	5.9	388	US-09-123-615-4	Sequence 4, Appl
8	113.5	5.1	418	US-08-855-910-11	Sequence 11, Appl
9	109	4.9	377	US-09-352-990-28	Sequence 28, Appl
10	108	4.8	197	US-08-923-867-2	Sequence 2, Appl
11	108	4.8	197	US-08-928-100-4	Sequence 4, Appl
12	108	4.8	197	US-09-183-134-2	Sequence 2, Appl
13	108	4.8	197	US-09-492-581-4	Sequence 4, Appl
14	104.5	4.7	370	US-08-415-593-45	Sequence 45, Appl
15	95	4.2	418	US-08-844-054-2	Sequence 2, Appl
16	95	4.2	418	US-09-347-333-2	Sequence 8, Appl
17	94.5	4.2	877	US-08-907-166-8	Sequence 8, Appl
18	90.5	4.0	898	US-08-465-995A-4	Sequence 4, Appl
19	90.5	4.0	898	US-08-465-994C-4	Sequence 4, Appl
20	90.5	4.0	898	US-08-966-145-4	Sequence 4, Appl
21	90.5	4.0	920	US-08-101-593-4	Sequence 4, Appl
22	89.5	4.0	344	US-09-393-554-2	Sequence 2, Appl
23	89.5	4.0	898	US-08-465-995A-2	Sequence 2, Appl
24	89.5	4.0	898	US-08-465-994C-2	Sequence 2, Appl
25	89.5	4.0	898	US-08-966-145-2	Sequence 2, Appl
26	89.5	4.0	920	US-08-101-593-2	Sequence 2, Appl
27	89	4.0	719	US-08-082-849B-31	Sequence 31, Appl

28	89	4.0	719	5	PCT-US94-01624-31	Sequence 31, Appl
29	88	3.9	606	2	US-08-883-534-3	Sequence 3, Appl
30	88	3.9	606	3	US-09-204-764-3	Sequence 3, Appl
31	87.5	3.9	855	4	US-08-880-865A-10	Sequence 10, Appl
32	87	3.9	428	3	US-08-331-625A-43	Sequence 43, Appl
33	87	3.9	1452	3	US-08-331-625A-2	Sequence 2, Appl
34	87	3.9	1452	5	PCT-US93-04384-18	Sequence 18, Appl
35	87	3.9	1452	5	PCT-US93-04692-2	Sequence 2, Appl
36	86	3.8	1443	1	US-08-308-872B-2	Sequence 2, Appl
37	85	3.8	1410	4	US-09-352-990-16	Sequence 16, Appl
38	83.5	3.7	510	1	US-08-249-112-3	Sequence 3, Appl
39	83.5	3.7	510	5	PCT-US95-06556-3	Sequence 3, Appl
40	83.5	3.7	970	1	US-08-375-709-7	Sequence 7, Appl
41	83.5	3.7	970	4	US-08-752-929-7	Sequence 7, Appl
42	83.5	3.7	970	4	US-09-090-793-5	Sequence 5, Appl
43	83	3.7	400	4	US-09-352-990-18	Sequence 18, Appl
44	83	3.7	551	1	US-08-484-493-15	Sequence 15, Appl
45	83	3.7	551	1	US-08-484-494-15	Sequence 15, Appl

## ALIGNMENTS

## RESULT 1

US-08-876-885-26  
Sequence 26, Application US/08876885

Patent No. 6174713  
GENERAL INFORMATION:

APPLICANT: Shen, Xiaoyu

APPLICANT: Houman, Fariba

TITLE OF INVENTION: CANDIDA CYTOPLASMIC TRYPTOPHANYL-tRNA

TITLE OF INVENTION: SYNTHETASE PROTEINS, NUCLEIC ACIDS AND STRAINS COMPRISING

TITLE OF INVENTION: SAME

NUMBER OF SEQUENCES: 26

CORRESPONDENCE ADDRESS:

ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.

STREET: Two Millitia Drive

CITY: Lexington

STATE: MA

COUNTRY: USA

ZIP: 02173

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/876,885

FILING DATE: 16-JUN-1997

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Brook, David E.

REGISTRATION NUMBER: 22,592

REFERENCE/DOCKET NUMBER: CPI97-02

TELECOMMUNICATION INFORMATION:

TELEPHONE: (781) 861-6240

TELEFAX: (781) 861-9540

INFORMATION FOR SEQ ID NO: 26:

SEQUENCE CHARACTERISTICS:

LENGTH: 424 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

US-08-876-885-26

## Query Match

Best Local Similarity 54.3%; Score 1218.5; DB 4; Length 424;  
Matches 226; Conservative 71; Mismatches 92; Indels 7; Gaps 3;

QY 31 TEAEEDVDPWTVO---TSSAKGIDYDKLIVRGSSKIDKELINRERATGCPHPHFLR 86  
Db 13 TESEOKITPEWVGAVVDCGSKMGIDYDKLISOPGTGTHTEETLERKQVYGEHPHFLK 72



CC uses, since the proteins are stable at very high temperatures, some up to  
 CC 110 degrees centigrade.  
 CC Note: This patent is in the same patent family as WO200065062, which  
 CC contains additional sequences as shown in AAB99143,  
 CC AAH75903-AAH75920 and AAG66436.

XX Sequence 385 AA;

Query Match 35.8%; Score 803; DB 22; Length 385;  
 Best Local Similarity 45.3%; Pred. No. 1.3e-73;  
 Matches 173; Conservative 63; Mismatches 120; Indels 26; Gaps 9;

QY 35 EDFVDPVOTSSAKGIDYDKLIVRGSSKIDKELINRIERATGQRPHEFLRGIFFSH 93  
 DB 3 EDFKVTWEVGV---VDYMKLIEHFGTSPLELELEKTAELTKSELPLFFRRKFFFSH 58  
 QY 94 RDMNOVLDAYENKPFYLYTGRSSSEAMHGHILPFTFTWLODVNPVLYIOMTDEK 153  
 DB 59 RDYDKVLQDYEEGRGFLLYGRGPSG--PMHGHILPFTATKWLQEKFGVNLXIQTDEK 117  
 QY 154 YLMD-LTLDQAYDAVENAKDIIACGFDIKFTIFSDLDYMGSSGFYKNVRIQKHT 212  
 DB 118 FLFENLTFEPTKHWATENIIDIAVGFDPKRTIFQNSER---TKYTEALFIKAKIN 173  
 QY 213 FNOVKGFIFGTFDSDIGKISFPATQAAFSNSPQIFRRTDIQCLIPCAIDDPYFRM 272  
 DB 174 FSKAKAVGFGTEQSKIGMIFPPAIQIAPTF-----FEKR---RCLIPAIQDDPTWRL 223  
 QY 273 TRDVAPRIGYKPLHLSTFPALQAGTQKMSASDPNSSILDTTAKOIKTKVKKHAFSG 332  
 DB 224 QRDFAESIGYKTAIHSKFVPSLTSQKMSASKPEATAYLTDSPEDVEKKVWFALTG 283  
 QY 333 GRDTEEHROBGNCDDVSPMYLTFPLEDDDKLEQIRKDY---TSGAMLTEGLKALIE 389  
 DB 284 GRPLKQREKGEPEKCVFKMLETFEEDDK--KLEKRYACKNGELTGCCKRILIS 341  
 QY 390 VLQPLIAHQARRKEVTDIYK 411  
 DB 342 KIQEFLKEHQRRKKAEKQIEK 363

RESULT 15  
 AAB58219  
 ID AAB58219 standard; Protein; 142 AA.

XX AAB58219;

DT 14-MAR-2001 (first entry)

DE Lung cancer associated polypeptide sequence SEQ ID 557.

XX Human; lung cancer associated protein; neuroprotective; cytostatic;  
 KW cardioactive; immunomodulatory; muscular active; vulnary;  
 KW gastrointestinal; nephrotropic; antiinfective; gynecological;  
 KW antibacterial; diagnosis; neural disorder; immune disorder; reproductive;  
 KW proliferative disorder; wound healing; infectious disease.

XX Homo sapiens.

XX WO200055180-A2.

PD 21-SEP-2000.

PF 08-MAR-2000; 2000MO-US05918.

XX 12-MAR-1999; 9905-0124270.

XX (HUMA-) HUMAN GENOME SCI INC.

PA (ROSE/) ROSEN C A.

XX Ruben SM;

XX WPI; 2000-587514/55.

DR N-PSDB: AAF18095.  
 XX Lung cancer associated gene sequences, referred to as lung cancer  
 PT antigens, useful for treatment, prevention, and diagnosis of disorders  
 PT such as lung cancer -

XX Claim 11, Page 1051; 1425pp; English.

XX Polynucleotide sequences AAF17982 - AAF18424 encode human lung cancer  
 CC associated proteins represented in AAB58106 - AAB58548. Lung cancer  
 CC associated proteins and polynucleotide sequences, their agonists, and  
 CC antagonists may have neuroprotective; cytostatic; cardioactive;  
 CC immunomodulatory; muscular active general; vulnary; gastrointestinal  
 CC general; nephrotropic; antiinfective; gynecological; or antibacterial  
 CC activity. The invention also includes antibodies specific for the  
 CC protein or polynucleotide sequences. The lung cancer associated  
 CC polynucleotide sequences may be used for detection of lung cancer,  
 CC chromosome identification, as chromosome markers, and for numerous other  
 CC diagnostic or research purposes. The proteins may be used to treat  
 CC disorders such as neural, immune, muscular, reproductive,  
 CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative  
 CC disorders. The proteins may also be used in the treatment of wounds and  
 CC infectious diseases. Polynucleotide sequences AAF18425 - AAF18433 and  
 CC peptide AAB58549 are used in the course of the invention for the  
 CC identification and characterisation of the polynucleotide and protein  
 CC sequences.

XX Sequence 142 AA;

Query Match 15.9%; Score 356.5; DB 21; Length 142;  
 Best Local Similarity 78.0%; Pred. No. 2.3e-28;  
 Matches 71; Conservative 5; Mismatches 12; Indels 3; Gaps 2;

QY 1 MSYAAAGEDYKADCPGPNAPTSNHPDATTEAEEDVDPVOTSSAKGIDYDKLIVRF 60  
 DB 52 MSYAAAGEDYKADCPGPNAPTSNHPDATTEAEEDVDPVOTSSAKGIDYDKLIVRF 111  
 QY 61 GSKIDKELINRIERATGQRPHEFLRGIF 91  
 DB 112 GSSXNQEELXDXP-STAKKTHS--GGGXFF 139

Search completed: October 24, 2002, 12:51:14  
 Job time : 27.5607 secs

PR 28-OCT-1999; 9905-0161993.  
PR 29-OCT-1999; 9905-0162142.

Query Match 43.1%; Score 968; DB 21; Length 292;  
Best Local Similarity 60.9%; Pred. No. 9e-91;  
Matches 179; Conservative 51; Mismatches 62; Indels 2; Gaps 1;

QY 131 IFTKMLQDVNVPVLIQMTDEKYLKMDLTLDQAYGAVENAKDIIACGPDINKTFIFSD 190  
DB 1 MFTKLGAFNVPVLIQLTDEKSLMKLSVESGRRLARENKADIIACGPDVTKTFIFSD 60  
QY 191 LDYMGSSGFFKKNVYKQKHVTENOVKGIQFTSDICIGKISPPAIQAAPFSNSFPQIF 250  
DB 61 FDYVG--GAFYKKNVYKVCYTLNKAMGIFGSGEDPIAKLSFPVQAVPSFSSFPHLF 118  
QY 251 RDRPTDIOCLIPCAIDODPYFRMTRDVAPRIGYKPPALHSFPFPAIQAGKMSADPNS 310  
DB 119 PGKDMRLCLIPCAIDODPYFRMTRDVAPRIGYKPPALHSFPFPAIQAGKMSADPNS 178  
QY 311 SIFLTPAKQIKTKVKNKHAFFSGGRDTIEHRQFGNCDVDVSFMYLTFLEDDDKLQIR 370  
DB 179 AIYVDSAKDIKKNINRYAFSGGDSIEKHRELKANLEVDIPVKYLSFLEDDSLLEHIK 238  
QY 371 KDYISGANLTGELKALLEVLOPLAEHQARKREYTDIYKEFMTPRKLSDFQ 424  
DB 239 KEYGGRMLTGEVKKRRLTEVLEIEKHHRRARAATAVDEMVAFAVAPRPLPSKFE 292

RESULT 13  
AAB58517  
ID AAB58517 standard; Protein; 173 AA.  
AC AAB58517;  
XX  
XX 14-MAR-2001 (first entry)  
DE Lung cancer associated polypeptide sequence SEQ ID 855.  
XX  
XX Human; lung cancer associated protein; neuroprotective; cytostatic;  
KW cardioactive; immunomodulatory; muscular active; vunerary;  
KW gastrointestinal; nephrotropic; antinefective; gynecological;  
KW antibacterial; diagnosis; neural disorder; immune disorder; reproductive;  
KW proliferative disorder; wound healing; infectious disease.  
XX  
XX Homo sapiens.  
OS  
XX  
XX MO20005180-A2.  
PN  
XX  
XX 21-SEP-2000.  
PD  
XX  
XX 08-MAR-2000; 2000MO-US05918.  
PE  
XX  
XX 12-MAR-1999; 99US-0124270.  
PR  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
PA  
XX  
XX (ROSE/) ROSEN C. A.  
PI  
XX  
XX Ruben SM;  
PI  
XX  
XX WPI: 2000-587514/55.  
DR N-PSDB; AAF18393.  
DR  
XX  
XX Lung cancer associated gene sequences, referred to as lung cancer  
PT antigens, useful for treatment, prevention, and diagnosis of disorders  
PT such as lung cancer -  
XX  
XX  
XX Claim 11; Page 1392-1393; 1425pp; English.  
XX  
XX Polynucleotide sequences AAF17982 - AAF18424 encode human lung cancer  
CC associated proteins represented in AAB58106 - AAB58548. Lung cancer  
CC associated proteins and polynucleotide sequences, their agonists, and  
CC antagonists may have neuroprotective; cytostatic; cardioactive;  
CC immunomodulatory; muscular active general; vunerary; gastrointestinal

CC general; nephrotropic; antinefective; gynecological; or antibacterial  
CC activity. The invention also includes antibodies specific for the  
CC protein or polynucleotide sequences. The lung cancer associated  
CC polynucleotide sequences may be used for detection of lung cancer,  
CC chromosome identification, as chromosome markers, and for numerous other  
CC diagnostic or research purposes. The proteins may be used to treat  
CC disorders such as neural, immune, muscular, reproductive,  
CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative  
CC disorders. The proteins may also be used in the treatment of wounds and  
CC infectious diseases. Polynucleotide sequences AAF18425 - AAF18433 and  
CC peptide AAB58549 are used in the course of the invention for the  
CC identification and characterisation of the polynucleotide and protein  
CC sequences.  
XX  
SQ Sequence 173 AA;  
Query Match 37.0%; Score 831; DB 21; Length 173;  
Best Local Similarity 98.7%; Pred. No. 5e-77;  
Matches 156; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 130 FIFTKMLQDVNVPVLIQMTDEKYLKMDLTLDQAYGAVENAKDIIACGPDINKTFIFS 189  
DB 1 FIFTKMLQDVNVPVLIQMTDEKYLKMDLTLDQAYGAVENAKDIIACGPDINKTFIFS 60  
QY 190 LDYMGSSGFFKKNVYKQKHVTENOVKGIQFTSDICIGKISPPAIQAAPFSNSFPQI 249  
DB 61 LDYMGSSGFFKKNVYKQKHVTENOVKGIQFTSDICIGKISPPAIQAAPFSNSFPQI 120  
QY 250 FRDRTDIOCLIPCAIDODPYFRMTRDVAPRIGYKPPAL 287  
DB 121 FRDRTDIOCLIPCAIDODPYFRMTRDVAPRIGYKPPAL 158

RESULT 14  
AAB96409  
ID AAB96409 standard; Protein; 385 AA.  
AC AAB96409;  
XX  
XX 29-OCT-2001 (first entry)  
DE Putative P. abyssi cryptophanyl-tRNA synthetase.  
XX  
XX Hyperthermophilic archaeon; hyperthermophilic protein.  
KW  
XX  
XX Pyrococcus abyssi.  
OS  
XX  
XX FR2792651-A1.  
PN  
XX  
XX 27-OCT-2000.  
PD  
XX  
XX 21-APR-1999; 99FR-0005034.  
PE  
XX  
XX 21-APR-1999; 99FR-0005034.  
PR  
XX  
XX (CNRS ) CNRS CENT NAT RECH SCI.  
PA (IFRE-) IFREMER INST FR RECH EXPL MER.  
XX  
XX  
XX Forterre P, Thierry JC, Prieur D, Dietrich J, Lecompte O;  
PI Querellou J, Weissenbach J, Saurin W, Helling R;  
PI  
XX  
XX WPI: 2001-126236/14.  
DR  
XX  
XX New nucleotide sequences isolated from Pyrococcus abyssi encode  
PT proteins useful in industry -  
PT  
XX  
XX  
XX Claim 7; Pages 1099-1100; 1657pp; French.  
XX  
XX The present invention relates to the genomic sequence of Pyrococcus  
CC abyssi (see AAF6441 and AAF11223-7) and P. abyssi proteins. P. abyssi is  
CC a hyperthermophilic archaeon, which is isolated from deep-sea  
CC hydrothermal vents. The present sequence is one such P. abyssi protein.  
CC The proteins of the present invention have various potential industrial

PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139452.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139763.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139889.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 02-JUL-1999; 99US-0142154.  
PR 06-JUL-1999; 99US-0142055.  
PR 08-JUL-1999; 99US-0142390.  
PR 09-JUL-1999; 99US-0142803.  
PR 12-JUL-1999; 99US-0142920.  
PR 13-JUL-1999; 99US-0142977.  
PR 14-JUL-1999; 99US-0143542.  
PR 15-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 21-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145218.  
PR 26-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 06-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 16-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149930.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 30-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151103.  
PR 31-AUG-1999; 99US-0151138.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158322.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160880.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.







```
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155139.
PR 24-SEP-1999; 99US-0155486.
PR 28-SEP-1999; 99US-0155659.
PR 29-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 26-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162142.
```

Query Match 58.1%; Score 1304; DB 21; Length 402;  
Best Local Similarity 60.3%; Pred. No. 4,1e-125;  
Matches 240; Conservative 67; Mismatches 87; Indels 4; Gaps 2;

```
QY 29 DATEE-EDVDPWTQYTSKAGIDYDKLYRFSSKIDKELINRIRARQGRHFLR 86
DB 7 DERESSEBOYVNPWEVSADGDKIDYDKLIDFGCORLDESLLDRVORLTSROPHVELR 66
QY 87 RGIFSRHDMNOVLDAVENKKPEYLYTGRGSSAMHGHILPFIETKMLDQVFWPLVI 146
DB 67 RSVFPAHNDPNEILDAYERGDKFYLYTGRGSSAMHGHILPFIETKMLDQVFWPLVI 126
QY 147 QMTDEKLYMKDLTDQAYGAVERNADIIACGFDINKTFIFSDIDYMGSSGFYKNVYK 206
DB 127 QLTDEKIMKLSVEESQRLARENAKDIIACGFDYTKTFIFSDIDYMG--GAFYKNVYK 184
QY 207 IQKHTFNQVNGIRFETSDCGKISFPAIQAPSFNSFQIFEDRDIDQICLPCALDQ 266
DB 185 VGCYTLKNAKIGFSESDPFAKLSFPVQAVPSFPLIFGKDNLCCLPCALDQ 244
QY 267 DPFYMTADVAPRIQYPPRIALHSTFPALQOATKMSASDPNSSIFLTDRAKQIKTKVN 326
DB 245 DPFYMTADVAPRIQYPPRIALHSTFPALQOATKMSASDPNSSIFLTDRAKQIKTKVN 304
QY 337 KHAESGDDTJEHRQFGNCDVDVSEMYLTFELEDKLEQIRKDYTSGMALGELKKA 386
DB 305 RYAESGGDSIEKRELANLEAVDIPVKYLSFELEDSELEIKREYEGHMLTGEVKKR 364
QY 387 LIEVLOPLIAHQARKREVTDEIYEFMTPRKLSFDQ 424
DB 365 LIEVLTIEVERHRRARAATVDEMADFAVVRPLPSKFE 402
```

RESULT 10  
AAG23697  
ID AAG23697 standard; Protein; 426 AA.  
XX  
AC AAG23697;  
XX  
DT 17-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 27100.  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
OS Arabidopsis thaliana.  
XX  
EN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PE 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.

PN EPI033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 9905-0121825.  
PR 05-MAR-1999; 9905-0123180.  
PR 09-MAR-1999; 9905-0123548.  
PR 23-MAR-1999; 9905-0125788.  
PR 25-MAR-1999; 9905-0126264.  
PR 29-MAR-1999; 9905-0126785.  
PR 01-APR-1999; 9905-0127462.  
PR 06-APR-1999; 9905-0128234.  
PR 08-APR-1999; 9905-0128714.  
PR 16-APR-1999; 9905-0129845.  
PR 19-APR-1999; 9905-0130077.  
PR 21-APR-1999; 9905-0130449.  
PR 23-APR-1999; 9905-0130510.  
PR 28-APR-1999; 9905-0130891.  
PR 30-APR-1999; 9905-0131449.  
PR 30-APR-1999; 9905-0132048.  
PR 04-MAY-1999; 9905-0132407.  
PR 05-MAY-1999; 9905-0132484.  
PR 06-MAY-1999; 9905-0132485.  
PR 06-MAY-1999; 9905-0132486.  
PR 07-MAY-1999; 9905-0132487.  
PR 11-MAY-1999; 9905-0132863.  
PR 14-MAY-1999; 9905-0134256.  
PR 14-MAY-1999; 9905-0134218.  
PR 14-MAY-1999; 9905-0134219.  
PR 14-MAY-1999; 9905-0134221.  
PR 18-MAY-1999; 9905-0134370.  
PR 19-MAY-1999; 9905-0134768.  
PR 20-MAY-1999; 9905-0134941.  
PR 21-MAY-1999; 9905-0135124.  
PR 25-MAY-1999; 9905-0135353.  
PR 25-MAY-1999; 9905-0135629.  
PR 27-MAY-1999; 9905-0136021.  
PR 28-MAY-1999; 9905-0136392.  
PR 01-JUN-1999; 9905-0136782.  
PR 03-JUN-1999; 9905-0137222.  
PR 04-JUN-1999; 9905-0137528.  
PR 07-JUN-1999; 9905-0137502.  
PR 08-JUN-1999; 9905-0137724.  
PR 10-JUN-1999; 9905-0138094.  
PR 10-JUN-1999; 9905-0138540.  
PR 14-JUN-1999; 9905-0138647.  
PR 16-JUN-1999; 9905-0139119.  
PR 16-JUN-1999; 9905-0139452.  
PR 17-JUN-1999; 9905-0139453.  
PR 17-JUN-1999; 9905-0139492.  
PR 18-JUN-1999; 9905-0139454.  
PR 18-JUN-1999; 9905-0139455.  
PR 18-JUN-1999; 9905-0139456.  
PR 18-JUN-1999; 9905-0139457.  
PR 18-JUN-1999; 9905-0139458.  
PR 18-JUN-1999; 9905-0139459.  
PR 18-JUN-1999; 9905-0139460.  
PR 18-JUN-1999; 9905-0139461.  
PR 18-JUN-1999; 9905-0139462.  
PR 18-JUN-1999; 9905-0139463.  
PR 18-JUN-1999; 9905-0139750.  
PR 18-JUN-1999; 9905-0139763.  
PR 21-JUN-1999; 9905-0139817.  
PR 22-JUN-1999; 9905-0139899.  
PR 23-JUN-1999; 9905-0140353.  
PR 23-JUN-1999; 9905-0140354.  
PR 24-JUN-1999; 9905-0140695.  
PR 28-JUN-1999; 9905-0140823.  
PR 29-JUN-1999; 9905-0140991.  
PR 30-JUN-1999; 9905-0141287.  
PR 01-JUL-1999; 9905-0141842.  
  
PR 01-JUL-1999; 9905-0142154.  
PR 02-JUL-1999; 9905-0142055.  
PR 06-JUL-1999; 9905-0142390.  
PR 08-JUL-1999; 9905-0142803.  
PR 09-JUL-1999; 9905-0142920.  
PR 12-JUL-1999; 9905-0142977.  
PR 13-JUL-1999; 9905-0143542.  
PR 14-JUL-1999; 9905-0143624.  
PR 15-JUL-1999; 9905-0144005.  
PR 16-JUL-1999; 9905-0144085.  
PR 16-JUL-1999; 9905-0144086.  
PR 19-JUL-1999; 9905-0144325.  
PR 19-JUL-1999; 9905-0144331.  
PR 19-JUL-1999; 9905-0144332.  
PR 19-JUL-1999; 9905-0144333.  
PR 19-JUL-1999; 9905-0144334.  
PR 19-JUL-1999; 9905-0144335.  
PR 20-JUL-1999; 9905-0144352.  
PR 20-JUL-1999; 9905-0144632.  
PR 20-JUL-1999; 9905-0144884.  
PR 21-JUL-1999; 9905-0144814.  
PR 21-JUL-1999; 9905-0145086.  
PR 21-JUL-1999; 9905-0145088.  
PR 22-JUL-1999; 9905-0145085.  
PR 22-JUL-1999; 9905-0145087.  
PR 22-JUL-1999; 9905-0145089.  
PR 22-JUL-1999; 9905-0145192.  
PR 23-JUL-1999; 9905-0145145.  
PR 23-JUL-1999; 9905-0145218.  
PR 23-JUL-1999; 9905-0145224.  
PR 26-JUL-1999; 9905-0145276.  
PR 27-JUL-1999; 9905-0145913.  
PR 27-JUL-1999; 9905-0145918.  
PR 27-JUL-1999; 9905-0145919.  
PR 28-JUL-1999; 9905-0145951.  
PR 02-AUG-1999; 9905-0146386.  
PR 02-AUG-1999; 9905-0146388.  
PR 02-AUG-1999; 9905-0146389.  
PR 03-AUG-1999; 9905-0147038.  
PR 04-AUG-1999; 9905-0147204.  
PR 04-AUG-1999; 9905-0147302.  
PR 05-AUG-1999; 9905-0147192.  
PR 05-AUG-1999; 9905-0147260.  
PR 06-AUG-1999; 9905-0147303.  
PR 06-AUG-1999; 9905-0147416.  
PR 09-AUG-1999; 9905-0147493.  
PR 09-AUG-1999; 9905-0147935.  
PR 10-AUG-1999; 9905-0148171.  
PR 11-AUG-1999; 9905-0148319.  
PR 12-AUG-1999; 9905-0148341.  
PR 13-AUG-1999; 9905-0148565.  
PR 13-AUG-1999; 9905-0148684.  
PR 16-AUG-1999; 9905-0149368.  
PR 17-AUG-1999; 9905-0149375.  
PR 18-AUG-1999; 9905-0149426.  
PR 20-AUG-1999; 9905-0149722.  
PR 20-AUG-1999; 9905-0149723.  
PR 20-AUG-1999; 9905-0149929.  
PR 23-AUG-1999; 9905-0149902.  
PR 23-AUG-1999; 9905-0149930.  
PR 26-AUG-1999; 9905-0150566.  
PR 26-AUG-1999; 9905-0150884.  
PR 27-AUG-1999; 9905-0151065.  
PR 27-AUG-1999; 9905-0151066.  
PR 27-AUG-1999; 9905-0151080.  
PR 30-AUG-1999; 9905-0151303.  
PR 31-AUG-1999; 9905-0151438.  
PR 01-SEP-1999; 9905-0151930.  
PR 07-SEP-1999; 9905-0152363.  
PR 10-SEP-1999; 9905-0153070.  
PR 13-SEP-1999; 9905-0153758.  
PR 15-SEP-1999; 9905-0154018.  
PR 16-SEP-1999; 9905-0154039.



QY 384 KKALEVLQPLIAEHQARRKEVTDEIVKEFMTPRKLSFDFQ 424  
 DB 362 KKALEVLQPLIAEHQARRKEVTDEIVKEFMTPRKLSFDFQ 402

RESULT 6  
 AAB47618 standard; Protein; 392 AA.  
 XX AAB47618:  
 AC AAB47618:  
 DT 07-JAN-2002 (first entry)  
 DE Human inactive TrpRS.  
 XX  
 KM Tyrosyl-tRNA synthetase; TyrRS; Rossmann fold nucleotide binding domain;  
 KM vascular endothelial cell function; burn; plastic surgery; abdomen;  
 KM polymorphonuclear leucocyte elastase; angiogenesis; tumor metastasis;  
 KM angiogenesis; graft; myocardial infarction; solid tumour; wound healing;  
 KM dermal ulcer; diabetic ulcer; endothelialization;  
 KM tryptophanyl-tRNA synthetase; trpRS; vascular graft surgery.  
 XX  
 OS Homo sapiens.  
 XX WO200174841-A1.  
 XX 11-OCT-2001.  
 PD 21-MAR-2001; 2001MO-US08966.  
 PF 31-MAR-2000; 2000US-193471P.  
 PR (SCRI ) SCRIPPS RES INST.  
 PA Schimmel P, Wakasugi K;  
 XX WPI: 2001-626377/72.  
 DR N-PSDB; AAB43605.  
 XX  
 PT New human truncated tyrosyl-tRNA synthetase polypeptide for regulating  
 PT vascular endothelial function, in particular for regulating  
 PT angiogenesis, tumor metastasis and treating myocardial infarction -  
 XX  
 PS Disclosure; Page 135-16; 150pp; English.  
 XX  
 CC The sequences given in AAB47615-18 show full length and truncated  
 CC versions of trptophanyl-tRNA synthetase (TrpRS). The truncated TrpRS of  
 CC the invention comprises a Rossmann fold nucleotide binding domain, and  
 CC is capable of regulating vascular endothelial cell function. It is of  
 CC approx. 40 kilo Dalton molecular weight and is produced by cleavage of  
 CC full length TrpRS with polymorphonuclear leucocyte elastase. Truncated  
 CC TrpRS is useful for regulating angiogenesis, tumor metastasis,  
 CC enhancing angiogenesis to a graft, treating myocardial infarction,  
 CC solid tumor, and a condition that would benefit from increased or  
 CC decreased angiogenesis in a mammal, in particular humans. It is also  
 CC useful in diagnosis and as a wound healing agent for treating wounds  
 CC such as dermal ulcers, diabetic ulcers, burns and injuries and in  
 CC plastic surgery when reconstruction is required following a burn or  
 CC for cosmetic purposes. It is particularly useful in the treatment of  
 CC abdominal wounds where there is high risk of infection. Truncated TrpRS  
 CC promotes endothelialization in vascular graft surgery and is used in  
 CC conjunction with angiography to administer the angiogenic tRNA  
 CC synthetase polypeptides or polynucleotides directly to the lumen and  
 CC wall of the blood vessel.  
 XX  
 SQ Sequence 392 AA:  
 Query Match 88.5%; Score 1988; DB 22; Length 392;  
 Best Local Similarity 100.0%; Pred. No. 1.8e-195.  
 Matches 378; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 47 SAKGIDYDKLIVRGSSKIDKELINRIERATGQRPNHFLRGJFFSHRDNQVLDAYENK 106  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

DB 2 SAKGIDYDKLIVRGSSKIDKELINRIERATGQRPNHFLRGJFFSHRDNQVLDAYENK 61  
 QY 107 KPEVLYTGRGSSPAMVGHILPITFTKWLQDVNVPLVITQMTDDEKYLKMDLTDQAYG 166  
 DB 62 KPEVLYTGRGSSPAMVGHILPITFTKWLQDVNVPLVITQMTDDEKYLKMDLTDQAYG 121  
 QY 167 DAVENAKDIACGFDINKTFFSDLDYMGSSGFGYKRVKVIQKHVTEPNQVKGIRGFTDSD 226  
 DB 122 DAVENAKDIACGFDINKTFFSDLDYMGSSGFGYKRVKVIQKHVTEPNQVKGIRGFTDSD 181  
 QY 227 CIGKISFPALQAPSEFNSFPQIFRDRTDIOCLIPCAIDDDPYRMRDYAPRIGYKPKA 286  
 DB 182 CIGKISFPALQAPSEFNSFPQIFRDRTDIOCLIPCAIDDDPYRMRDYAPRIGYKPKA 241  
 QY 287 LHSFPPALQAGTQKASASPNSIFLTPAKOIKTKVKNHAFSGGRDTIEEHROFGN 346  
 DB 242 LHSFPPALQAGTQKASASPNSIFLTPAKOIKTKVKNHAFSGGRDTIEEHROFGN 301  
 QY 347 CDVDVSEFMYLTFELEDDEKLEQIRKDYTSGAMLTGELKKALEVLQPLIAEHQARRKEVT 406  
 DB 302 CDVDVSEFMYLTFELEDDEKLEQIRKDYTSGAMLTGELKKALEVLQPLIAEHQARRKEVT 361  
 QY 407 DEIVKEFMTPRKLSFDFQ 424  
 DB 362 DEIVKEFMTPRKLSFDFQ 379

RESULT 7  
 ABB64621 standard; Protein; 430 AA.  
 XX ABB64621:  
 AC ABB64621:  
 DT 26-MAR-2002 (first entry)  
 DE Drosophila melanogaster polypeptide SEQ ID NO 20655.  
 XX  
 KM Drosophila: developmental biology; cell signalling; insecticide;  
 KM pharmaceutical.  
 KM Drosophila melanogaster.  
 XX  
 PD WO200171042-A2.  
 PF 27-SEP-2001.  
 PR 23-MAR-2001; 2001MO-US09231.  
 PR 23-MAR-2000; 2000US-191637P.  
 PR 11-JUL-2000; 2000US-0614150.  
 PA (PEKE ) PE CORP NY.  
 XX  
 XX Venter JC, Adams M, Li PWD, Myers EW;  
 PI WPI: 2001-656860/75.  
 DR N-PSDB; ABL08724.  
 XX  
 PT New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signalling and cell-cell  
 PT interactions -  
 XX  
 PS Disclosure; SEQ ID NO 20655; 21pp + Sequence Listing; English.  
 XX  
 CC The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
 CC sequences (ABL01840-ABL16175) and the encoded proteins  
 CC (ABBS7737-ABBS72072).  
 CC The sequence data for this patent did not form part of the printed

DR N-PSDB: AAX33942.  
XX  
PT New isolated human genes  
XX  
PS Claim 3; Page 112-114; 184pp; English.  
XX  
CC This sequence is encoded by a human gene of the invention, and is induced  
CC to express by both HCMV and interferon (IFN), designated HCMV-inducible  
CC genes (cig or cigs). The invention also relates to genes that are  
CC repressed in the presence of HCMV infection, designated HCMV-repressible  
CC genes (crg or crgs). The products can be used to obtain agents which can  
CC be used for anti-viral therapy, particularly anti-HCMV therapy. They can  
CC also be used for the development of drugs that would allow for higher  
CC dosage IFN treatments without the concomitant toxicity normally  
CC associated with administering high levels of IFN. The products can also  
CC be used for detection, diagnosis and drug screening.  
XX  
SQ Sequence 471 AA;  
Query Match 99.1%; Score 2226; DB 20; Length 471;  
Best Local Similarity 99.3%; Pred. No. 7.9e-220;  
Matches 421; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 1 MSYKAGEDYKADCPGPNAPTNNHGPDAEEDFVDPVTQTSKAGIDYDKLIVRF 60  
DB 48 MSYKAAAGEDYKADCPGPNAPTNNHGPDAEEDFVDPVTQTSKAGIDYDKLIVRF 107  
QY 61 GSSKIDKELINIERATGQRPHFRGFFFSHRDMNOVLAYENKKPFYLYTGRGPSE 120  
DB 108 GSSKIDKELINIERATGQRPHFRGFFFSHRDMNOVLAYENKKPFYLYTGRGPSE 167  
QY 121 AMHVGHLIPFTFKMLQDVFNPLVIOGMTDEKYLMDLTLDOAYGDAVENAKDIIACGF 180  
DB 168 AMHVGHLIPFTFKMLQDVFNPLVIOGMTDEKYLMDLTLDOAYGDAVENAKDIIACGF 227  
QY 181 DINKTFITSDLDYMGSSGFKYNNVKIQKHVTFNOYKGIFFGTTSDCIIGKISFPAIQAP 240  
DB 228 DINKTFITSDLDYMGSSGFKYNNVKIQKHVTFNOYKGIFFGTTSDCIIGKISFPAIQAP 287  
QY 241 SFSNFPQIFRRDIOCLIPCAIDODPFRRMTRDVAPRIGYPRKALHSTFFPALQAGQ 300  
DB 288 SFSNFPQIFRRDIOCLIPCAIDODPFRRMTRDVAPRIGYPRKALHSTFFPALQAGQ 347  
QY 301 TKMSASDPNSSIFLTDPAKQIKTKVKNHAFSGGRDIEEHROFGNCVDVYSFMYLTFFFL 360  
DB 348 TKMSASDPNSSIFLTDPAKQIKTKVKNHAFSGGRDIEEHROFGNCVDVYSFMYLTFFFL 407  
QY 361 EDDDKLEQIRKDYTGAMLTGELKALIEVLIQPLIAEHQARRKEVTDIYKEFMTPRKLS 420  
DB 408 EDDDKLEQIRKDYTGAMLTGELKALIEVLIQPLIAEHQARRKEVTDIYKEFMTPRKLS 467  
QY 421 FDFQ 424  
DB 468 FDFQ 471

RESULT 5  
AAB47617  
ID AAB47617 standard; Protein: 415 AA.

XX AAB47617;

DT 07-JAN-2002 (first entry)

DE Human supermini TrpRS.

XX Tyrosyl-tRNA synthetase; TrpRS; Rossmann fold nucleotide binding domain;  
KW vascular endothelial cell function; burn; plastic surgery; abdomen;  
KW polymorphonuclear leucocyte elastase; angiogenesis; tumour metastasis;  
KW angiogenesis; graft; myocardial infarction; solid tumour; wound healing;  
KW dermal ulcer; diabetic ulcer; endothelialization;  
KW tryptophanyl-tRNA synthetase; trpRS; vascular graft surgery.  
XX

OS Homo sapiens.  
XX  
PN WO200174841-A1.  
XX  
XX 11-OCT-2001.  
PD  
XX 21-MAR-2001; 2001WO-US08966.  
PF  
XX 31-MAR-2000; 2000US-193471P.  
PR  
XX (SCRI ) SCRIPPS RES INST.  
PA  
XX Schimmel P, Wakasugi K;  
PI  
XX WPI: 2001-626377/72.  
DR  
XX N-PSDB: AAB43604.  
PT New human truncated tyrosyl-tRNA synthetase polypeptide for regulating  
PT vascular endothelial function, in particular for regulating  
PT angiogenesis, tumor metastasis and treating myocardial infarction -  
PS Disclosure; Page 129-30; 150pp; English.  
XX  
CC The sequences given in AAB47615-18 show full length and truncated  
CC versions of tryptophanyl-tRNA synthetase (TrpRS). The truncated TrpRS of  
CC the invention comprises a Rossmann fold nucleotide binding domain, and  
CC is capable of regulating vascular endothelial cell function. It is of  
CC approx. 40 kDa molecular weight and is produced by cleavage of  
CC full length TrpRS with polymorphonuclear leucocyte elastase. Truncated  
CC TrpRS is useful for regulating angiogenesis, tumor metastasis,  
CC enhancing angiogenesis to a graft, treating myocardial infarction,  
CC solid tumor, and a condition that would benefit from increased or  
CC decreased angiogenesis in a mammal, in particular humans. It is also  
CC useful in diagnosis and as a wound healing agent for treating wounds  
CC such as dermal ulcers, diabetic ulcers, burns and injuries and in  
CC plastic surgery when reconstruction is required following a burn or  
CC for cosmetic purposes. It is particularly useful in the treatment of  
CC abdominal wounds where there is high risk of infection. Truncated TrpRS  
CC promotes endothelialization in vascular graft surgery and is used in  
CC conjunction with angiography to administer the angiogenic RNA  
CC synthetase polypeptides or polynucleotides directly to the lumen and  
CC wall of the blood vessel.  
XX  
SQ Sequence 415 AA;  
Query Match 94.2%; Score 2116; DB 22; Length 415;  
Best Local Similarity 100.0%; Pred. No. 1.3e-208;  
Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 24 SNHGPDATAEEDFVDPVTQTSKAGIDYDKLIVRFSSKIDKELINIERATGQRPHN 83  
DB 2 SNHGPDATAEEDFVDPVTQTSKAGIDYDKLIVRFSSKIDKELINIERATGQRPHN 61  
QY 84 FLRGIFFSHRDMNOVLDAYENKKPFYLYTGRGPSSAMHVGHLIPFTFKMLQDVFNVP 143  
DB 62 FLRGIFFSHRDMNOVLDAYENKKPFYLYTGRGPSSAMHVGHLIPFTFKMLQDVFNVP 121  
QY 144 LVIOGMTDEKYLMDLTLDOAYGDAVENAKDIIACGPIINTFTFSDLDYMGSSGFKYNN 203  
DB 122 LVIOGMTDEKYLMDLTLDOAYGDAVENAKDIIACGPIINTFTFSDLDYMGSSGFKYNN 181  
QY 204 VKIQRKHYTFNOYKGIFFGTTSDCIIGKISFPAIQAPFSNSFPQIFRDRDIOCLIPCA 263  
DB 182 VKIQRKHYTFNOYKGIFFGTTSDCIIGKISFPAIQAPFSNSFPQIFRDRDIOCLIPCA 241  
QY 264 IDDDPYFRMTRDVAPRIGYPRKALHSTFFPALQAGQTKMSASDPNSSIFLTDPAKQIKT 323  
DB 242 IDDDPYFRMTRDVAPRIGYPRKALHSTFFPALQAGQTKMSASDPNSSIFLTDPAKQIKT 301  
QY 324 KVNKHAFFSGGRDIEEHROFGNCVDVYSFMYLTFFLEDDDKLEQIRKDYTGAMLTGEL 383  
DB 302 KVNKHAFFSGGRDIEEHROFGNCVDVYSFMYLTFFLEDDDKLEQIRKDYTGAMLTGEL 361

QY 241 SFSNSFPIFRDRDIOCLIPCAIDODPYFRMTDVAIRIGYPPKALLHSTFFPALGQAQ 300  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 288 SFSNSFPIFRDRDIOCLIPCAIDODPYFRMTDVAIRIGYPPKALLHSTFFPALGQAQ 347  
 QY 301 TKMSASDPNSSIFLDTAKQIKTKVKNKHAFFSGRDTIEHROFGNCDVDVSFMYLTFEFL 360  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 348 TKMSASDPNSSIFLDTAKQIKTKVKNKHAFFSGRDTIEHROFGNCDVDVSFMYLTFEFL 407  
 QY 361 EDDDKLEQIRKDYTSGAMLTGELKKALIEVQLPIAEHQARKKEVTDEIVAEFMTPKRLS 420  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 408 EDDDKLEQIRKDYTSGAMLTGELKKALIEVQLPIAEHQARKKEVTDEIVAEFMTPKRLS 467  
 QY 421 FDFQ 424  
 ||||  
 Db 468 FDFQ 471

RESULT 3  
 AAB58220  
 ID AAB58220 standard; Protein: 475 AA.  
 XX  
 AC AAB58220;  
 XX  
 DT 14-MAR-2001 (first entry)  
 XX  
 DE Lung cancer associated polypeptide sequence SFQ ID 558.  
 XX  
 KW Human; lung cancer associated protein; neuroprotective; cytoskeletal;  
 KM cardioactive; immunomodulatory; muscular active; vulnery;  
 KM gastrointestinal; nephrotropic; antiinfective; gynecological;  
 KM antibacterial; diagnosis; neural disorder; immune disorder; reproductive;  
 KM proliferative disorder; wound healing; infectious disease.  
 XX  
 OS Homo sapiens.  
 XX  
 PN W0200055180-A2.  
 XX  
 PD 21-SEP-2000.  
 XX  
 PF 08-MAR-2000; 2000WO-US05918.  
 XX  
 PR 12-MAR-1999; 99US-0124270.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 PA (ROSE/) ROSEN C A.  
 XX  
 PI Ruben SM;  
 XX  
 DR WPI: 2000-587514/55.  
 DR N-PSDB: AAF18096.  
 XX  
 PT Lung cancer associated gene sequences, referred to as lung cancer  
 PT antigens, useful for treatment, prevention, and diagnosis of disorders  
 PT such as lung cancer -  
 XX  
 PS Claim 11; Page 1052-1053; 1425pp; English.  
 XX  
 CC Polynucleotide sequences AAF17982 - AAF18424 encode human lung cancer  
 CC associated proteins represented in AAB58106 - AAB58548. Lung cancer  
 CC associated proteins and polynucleotide sequences, their agonists, and  
 CC antagonists may have neuroprotective; cytoskeletal; cardioactive;  
 CC immunomodulatory; muscular active general; vulnery; gastrointestinal  
 CC general; nephrotropic; antiinfective; gynecological; or antibacterial  
 CC activity. The invention also includes antibodies specific for the  
 CC protein or polynucleotide sequences. The lung cancer associated  
 CC polynucleotide sequences may be used for detection of lung cancer,  
 CC chromosome identification, as chromosome markers, and for numerous other  
 CC diagnostic or research purposes. The proteins may be used to treat  
 CC disorders such as neural, immune, muscular, reproductive,  
 CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative  
 CC disorders. The proteins may also be used in the treatment of wounds and  
 CC infectious diseases. Polynucleotide sequences AAF18425 - AAF18433 and

CC peptide AAB58549 are used in the course of the invention for the  
 CC identification and characterisation of the polynucleotide and protein  
 CC sequences.  
 XX  
 XX Sequence 475 AA;  
 SQ  
 Query Match 99.3%; Score 2231; DB 21; Length 475;  
 Best Local Similarity 99.5%; Pred. No. 2,4e-220;  
 Matches 422; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 MSYKAAGEDYKADCPGPNPAFTSNHGPDATEAEEDFVDPTVOTSSAKGIDYDKLYRF 60  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 52 MSYKAAGEDYKADCPGPNPAFTSNHGPDATEAEEDFVDPTVOTSSAKGIDYDKLYRF 111  
 QY 61 GSSKIDKELNRIERATGQRPHHFLRGIFPSHRDMNOVLDAVENKKPEVLYTGRPSSE 120  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 112 GSSKIDKELNRIERATGQRPHHFLRGIFPSHRDMNOVLDAVENKKPEVLYTGRPSSE 171  
 QY 121 AMHYGHLIPFIETKMLDVFNVPLVIQMTDDEKYLKDLTDQAYGDAVENAKDIACGF 180  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 172 AMHYGHLIPFIETKMLDVFNVPLVIQMTDDEKYLKDLTDQAYGDAVENAKDIACGF 231  
 QY 181 DINKTFIFSDLDYMGMSGGFYKNVYKIQKHVTENQVKGIFGFTSDSICIGISFPALQAP 240  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 232 DINKTFIFSDLDYMGMSGGFYKNVYKIQKHVTENQVKGIFGFTSDSICIGISFPALQAP 291  
 QY 241 SFSNSFPIFRDRDIOCLIPCAIDODPYFRMTDVAIRIGYPPKALLHSTFFPALGQAQ 300  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 292 SFSNSFPIFRDRDIOCLIPCAIDODPYFRMTDVAIRIGYPPKALLHSTFFPALGQAQ 351  
 QY 301 TKMSASDPNSSIFLDTAKQIKTKVKNKHAFFSGRDTIEHROFGNCDVDVSFMYLTFEFL 360  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 352 TKMSASDPNSSIFLDTAKQIKTKVKNKHAFFSGRDTIEHROFGNCDVDVSFMYLTFEFL 411  
 QY 361 EDDDKLEQIRKDYTSGAMLTGELKKALIEVQLPIAEHQARKKEVTDEIVAEFMTPKRLS 420  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 412 EDDDKLEQIRKDYTSGAMLTGELKKALIEVQLPIAEHQARKKEVTDEIVAEFMTPKRLS 471  
 QY 421 FDFQ 424  
 ||||  
 Db 472 FDFQ 475

RESULT 4  
 AAY05372  
 ID AAY05372 standard; Protein: 471 AA.  
 XX  
 AC AAY05372;  
 XX  
 DT 30-JUN-1999 (first entry)  
 XX  
 DE Human HCMV inducible gene protein, SEQ ID NO 12.  
 XX  
 KW HCMV inducible gene; cig; human; human cytomegalovirus; interferon;  
 KW anti-viral therapy; anti-HCMV therapy; detection; diagnosis;  
 KW drug screening.  
 XX  
 OS Homo sapiens.  
 XX  
 PN W09913075-A2.  
 XX  
 PD 18-MAR-1999.  
 XX  
 PF 08-SEP-1998; 98WO-US18638.  
 XX  
 PR 22-SEP-1997; 97US-0059725.  
 PR 08-SEP-1997; 97US-0058180.  
 XX  
 PA (UTPR-) UNIV PRINCETON.  
 XX  
 PI Cong J, Schenk T, Zhu H;  
 XX  
 DR WPI: 1999-243729/20.

PT New human truncated tyrosyl-tRNA synthetase polypeptide for regulating  
PT vascular endothelial function, in particular for regulating  
PT angiogenesis, tumor metastasis and treating myocardial infarction -  
XX  
PS Disclosure: Page 123-24; 150pp; English.  
XX  
CC The sequences given in AAB47615-18 show full length and truncated  
CC versions of tryptophanyl-tRNA synthetase (TrpRS). The truncated TrpRS of  
CC the invention comprises a Rossmann fold nucleotide binding domain, and  
CC is capable of regulating vascular endothelial cell function. It is of  
CC approx. 40 kDa Dalton molecular weight and is produced by cleavage of  
CC full length TrpRS with polynucleonuclear leucocyte elastase. Truncated  
CC TrpRS is useful for regulating angiogenesis, tumor metastasis,  
CC enhancing angiogenesis to a graft, treating myocardial infarction,  
CC solid tumor, and a condition that would benefit from increased or  
CC decreased angiogenesis in a mammal, in particular humans. It is also  
CC useful in diagnosis and as a wound healing agent for treating wounds  
CC such as dermal ulcers, diabetic ulcers, burns and injuries and in  
CC plastic surgery when reconstruction is required following a burn or  
CC for cosmetic purposes. It is particularly useful in the treatment of  
CC abdominal wounds where there is high risk of infection. Truncated TrpRS  
CC promotes endothelialization in vascular graft surgery and is used in  
CC conjunction with angiography to administer the angiogenic tRNA  
CC synthetase polypeptides or polynucleotides directly to the lumen and  
CC wall of the blood vessel.  
XX  
XX

Sequence 437 AA;

Query Match 100.0%; Score 2246; DB 22; Length 437;  
Best Local Similarity 100.0%; Pred. No. 6.2e-222;  
Matches 424; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSYRAAGEDYKACDPCPNPAPTNSHGPDATEAEDPDPWTVOTSSAKGIDYKLYRF 60  
DB 1 MSYRAAGEDYKACDPCPNPAPTNSHGPDATEAEDPDPWTVOTSSAKGIDYKLYRF 60  
QY 61 GSSKIDKELINRIERATGQRPHHFLRGIFFSHRDMNOVLDAENKKPFYLYTGRGSSSE 120  
DB 61 GSSKIDKELINRIERATGQRPHHFLRGIFFSHRDMNOVLDAENKKPFYLYTGRGSSSE 120  
QY 121 AMHVGHLIPFTFKWLODVFNVPVLTQMTDDEKYLKMDLTLDQAYGDAVENAKDIIACGF 180  
DB 121 AMHVGHLIPFTFKWLODVFNVPVLTQMTDDEKYLKMDLTLDQAYGDAVENAKDIIACGF 180  
QY 181 DINKTFISDDIDYMGSSGCFKYNVVKIOKHTFNQVKGIFGFTSDCIGKISFPAIQAAP 240  
DB 181 DINKTFISDDIDYMGSSGCFKYNVVKIOKHTFNQVKGIFGFTSDCIGKISFPAIQAAP 240  
QY 241 SFSNSFPQIFRDRDIOCLIPCAIDQDPPYFRMTRDVABRIGYKPPALLHSTFFPALQGAQ 300  
DB 241 SFSNSFPQIFRDRDIOCLIPCAIDQDPPYFRMTRDVABRIGYKPPALLHSTFFPALQGAQ 300  
QY 301 TKMSASDNNSSIFLTDYAKQIKTKYKNAHFGSGRDTIEHNOFGNCVDVSMYLTFFL 360  
DB 301 TKMSASDNNSSIFLTDYAKQIKTKYKNAHFGSGRDTIEHNOFGNCVDVSMYLTFFL 360  
QY 361 EDDDLLEGIRKDYTSGAULTGELKKALIEVLOPLAEHQARRKEVTDIVKEFMPRKLS 420  
DB 361 EDDDLLEGIRKDYTSGAULTGELKKALIEVLOPLAEHQARRKEVTDIVKEFMPRKLS 420  
QY 421 FDFQ 424  
DB 421 FDFQ 424

RESULT 2  
AAB47615  
ID AAB47615 standard; Protein: 484 AA.

AC AAB47615;

XX 07-JAN-2002 (first entry)

XX

DE Human full-length TrpRS.

XX Tyrosyl-tRNA synthetase; TyRS; Rossmann fold nucleotide binding domain;  
KW vascular endothelial cell function; burn; plastic surgery; abdomen;  
KW polymorphonuclear leucocyte elastase; angiogenesis; tumour metastasis;  
KW angiogenesis; graft; myocardial infarction; solid tumour; wound healing;  
KW dermal ulcer; diabetic ulcer; endothelialization;  
KW tryptophanyl-tRNA synthetase; trpRS; vascular graft surgery.  
XX

OS Homo sapiens.

PN W0200174841-A1.

XX 11-OCT-2001.

PF 21-MAR-2001; 2001MO-US080966.

PR 31-MAR-2000; 2000US-193471P.

XX (SCRI ) SCRIPPS RES INST.

PI Schimmel P, Wakasugi K;

XX WPI: 2001-626377/72.

DR N-PSDB; AAB43602.

XX

PT New human truncated tyrosyl-tRNA synthetase polypeptide for regulating

PT vascular endothelial function, in particular for regulating

PT angiogenesis, tumor metastasis and treating myocardial infarction -

XX  
PS Disclosure: Page 117-19; 150pp; English.

CC The sequences given in AAB47615-18 show full length and truncated  
CC versions of tryptophanyl-tRNA synthetase (TrpRS). The truncated TrpRS of  
CC the invention comprises a Rossmann fold nucleotide binding domain, and  
CC is capable of regulating vascular endothelial cell function. It is of  
CC approx. 40 kDa Dalton molecular weight and is produced by cleavage of  
CC full length TrpRS with polynucleonuclear leucocyte elastase. Truncated  
CC TrpRS is useful for regulating angiogenesis, tumor metastasis,  
CC enhancing angiogenesis to a graft, treating myocardial infarction,  
CC solid tumor, and a condition that would benefit from increased or  
CC decreased angiogenesis in a mammal, in particular humans. It is also  
CC useful in diagnosis and as a wound healing agent for treating wounds  
CC such as dermal ulcers, diabetic ulcers, burns and injuries and in  
CC plastic surgery when reconstruction is required following a burn or  
CC for cosmetic purposes. It is particularly useful in the treatment of  
CC abdominal wounds where there is high risk of infection. Truncated TrpRS  
CC promotes endothelialization in vascular graft surgery and is used in  
CC conjunction with angiography to administer the angiogenic tRNA  
CC synthetase polypeptides or polynucleotides directly to the lumen and  
CC wall of the blood vessel.  
XX

Sequence 484 AA;

Query Match 100.0%; Score 2246; DB 22; Length 484;  
Best Local Similarity 100.0%; Pred. No. 7.2e-222;  
Matches 424; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSYRAAGEDYKACDPCPNPAPTNSHGPDATEAEDPDPWTVOTSSAKGIDYKLYRF 60  
DB 48 MSYRAAGEDYKACDPCPNPAPTNSHGPDATEAEDPDPWTVOTSSAKGIDYKLYRF 107  
QY 61 GSSKIDKELINRIERATGQRPHHFLRGIFFSHRDMNOVLDAENKKPFYLYTGRGSSSE 120  
DB 108 GSSKIDKELINRIERATGQRPHHFLRGIFFSHRDMNOVLDAENKKPFYLYTGRGSSSE 167  
QY 121 AMHVGHLIPFTFKWLODVFNVPVLTQMTDDEKYLKMDLTLDQAYGDAVENAKDIIACGF 180  
DB 168 AMHVGHLIPFTFKWLODVFNVPVLTQMTDDEKYLKMDLTLDQAYGDAVENAKDIIACGF 227  
QY 181 DINKTFISDDIDYMGSSGCFKYNVVKIOKHTFNQVKGIFGFTSDCIGKISFPAIQAAP 240  
DB 228 DINKTFISDDIDYMGSSGCFKYNVVKIOKHTFNQVKGIFGFTSDCIGKISFPAIQAAP 287

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 24, 2002, 12:49:40 ; Search time 26.5607 Seconds  
(without alignments)  
1773.117 Million cell updates/sec

Title: US-09-813-718-10\_COPY\_48\_471

Perfect score: 2246  
Sequence: 1 MSYKAAGCEDYKADCPGPNF.....VTDEIVKEFMTPRKLSDFDQ 424

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A\_Geneseq\_032802:.\*  
1: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2246	100.0	437	22	AA847616 Human mini TyPRS.
2	2246	100.0	484	22	AA847615 Human full-length
3	2231	99.3	475	21	AA858220 Lung cancer associ
4	2226	99.1	471	20	AAV05372 Human HCMV induci
5	2116	94.2	415	22	AA847617 Human supermini Tr
6	1988	88.5	392	22	AA847618 Human inactive Trp
7	1378	61.4	430	22	ABB64621 Drosophila melanog
8	1378	61.4	430	22	ABB67203 Drosophila melanog
9	1304	58.1	426	21	AA833698 Arabidopsis thalia
10	1304	58.1	426	21	AA833697 Arabidopsis thalia
11	1218.5	54.3	424	22	AA866931 Tryptophanyl-LRNA

12	968	43.1	292	21	AA833699 Arabidopsis thalia
13	831	37.0	173	21	AA858517 Lung cancer associ
14	803	35.8	385	22	AA896409 Putative P. abyss
15	356.5	15.9	142	21	AA858219 Lung cancer associ
16	185.5	8.3	341	19	AA856423 Tryptophanyl-LRNA
17	185.5	8.3	341	22	AA837851 Streptococcus pneu
18	164.5	7.3	409	20	AA894248 C. albicans tyrosy
19	162.5	7.2	409	20	AA894247 C. albicans tyrosy
20	159.5	7.1	379	20	AA837623 Protein involved i
21	158.5	7.1	378	22	AA896600 Putative P. abyss
22	152	6.8	344	20	AA835439 Chlamydia pneumoni
23	146	6.5	339	22	AA835829 Helicobacter pylori
24	138.5	6.2	140	21	AA833032 Arabidopsis thalia
25	134.5	6.0	525	22	AA860745 Drosophila melanog
26	134	6.0	372	22	AA847614 Human mini TyPRS m
27	132.5	5.9	372	22	AA847612 Human mini TyPRS
28	132.5	5.9	536	22	AA847611 Human full length
29	128	5.7	347	21	AA835477 Haemophilus influe
30	123	5.5	347	21	AA831904 Arabidopsis thalia
31	122.5	5.5	346	21	AA848617 Arabidopsis thalia
32	120.5	5.4	404	22	AA854020 Propionibacterium
33	115.5	5.1	129	21	AA833033 Arabidopsis thalia
34	113.5	5.1	418	22	AA802578 Enterococcus tyros
35	113.5	5.1	423	22	AA835085 Enterococcus faeca
36	113	5.0	415	22	AA833998 Staphylococcus aur
37	113	5.0	420	18	AA819781 Tyrosyl-tRNA synth
38	113	5.0	420	22	AA837056 Staphylococcus aur
39	112	5.0	345	22	AA890498 C glutamicum prote
40	108	4.8	344	22	AA838213 Salmonella typhi c
41	107	4.8	71	22	AA810441 Human polypeptide
42	106.5	4.7	334	22	AA834778 E. coli cellular p
43	104.5	4.7	370	20	AA817508 pneumocystis carin
44	100.5	4.5	421	22	AA882363 S. epidermidis ope
45	95	4.2	418	18	AA827663 Streptococcus pneu

# ALIGNMENTS

RESULT 1	AA847616 standard; Protein: 437 AA.
ID	AA847616
XX	XX
AC	AA847616;
XX	XX
DT	07-JAN-2002 (first entry)
XX	XX
DE	Human mini TyPRS.
XX	XX
KW	Tyrosyl-tRNA synthetase; TyPRS: Rossmann fold nucleotide binding domain;
KW	vascular endothelial cell function; burn; plastic surgery; tumour metastasis;
KW	polymorphonuclear leucocyte elastase; angiogenesis; tumour metastasis;
KW	angiogenesis; graft; myocardial infarction; solid tumour; wound healing;
KW	dermal ulcer; endothelialization;
KW	tryptophanyl-tRNA synthetase; tyPRS; vascular graft surgery.
XX	XX
OS	Homo sapiens.
XX	XX
PN	WO200174841-A1.
XX	XX
PD	11-OCT-2001.
XX	XX
PF	21-MAR-2001; 2001MO-US08966.
XX	XX
PR	31-MAR-2000; 2000US-193471P.
XX	XX
PA	(SCRI ) SCRIIPS RES INST.
XX	XX
PI	Schimmel P, Wakasugi K;
XX	XX
DR	WPI; 2001-626377/2.
XX	XX
DR	N-PSDB; AA843603.
XX	XX



RL DNA Res. 5:55-76(1998).  
 DR EMBL: AP000007; BAA31046.1; .  
 DR InterPro: IPR002305; tRNA-synt\_1b.  
 DR InterPro: IPR002306; tRNA-synt\_trp.  
 DR Pfam: PF00579; tRNA-synt\_1b; 1.  
 DR PRINTS: PR01039; TRNASYNTHTRP.  
 KW Aminoacyl-tRNA synthetase; Complete proteome.  
 SQ SEQUENCE 301 AA; 35312 MW; 957143BE1647CAE1 CRC64;  
  
 Query Match 24.6%; Score 626.5; DB 17; Length 301;  
 Best Local Similarity 44.2%; Pred. No. 2e-45; Mismatches 92; Indels 21; Gaps 7;  
 Matches 134; Conservative 56; Mismatches 92; Indels 21; Gaps 7;  
  
 QY 169 MHVGLIPETITKMLQDVENVPLVIOMTDEKYLKMD-LTLDOAYGDAVENAKDIACGF 227  
 DB 1 MHGHTIPFATKMLQEKGVNLVYIOTDDEKFLFRENLTFFDTRKMAVDNLIDIAVGF 60  
 QY 228 DINKFTFESDLVGMSSGFYKNVYKIQKHVFNVOYKGIFFGTDSDCIKISFPALQAP 287  
 DB 61 DDDKFTFIFQNSF---TKIYEMAIPIAKKINFSMAKAVFGFTSQIKGIMIFFPAIQIAP 116  
 QY 288 SFSNSFPQIFRDRTDIOCLIPCAIDODPYFRMTRDVAPRIGYKPKALLHSTFEPALQAG 347  
 DB 117 TF-----FEKK---KCLIPALIDDDPYWRLORDPAESLGYYKTRALHSKFPVSLTSL 166  
 QY 348 TKMSASDPNSSIFLDTAKQIKTKYKNAFSGGRDTEIEHRQFGNCDDVDSFMYLTFPL 407  
 DB 167 GKMSASKEPTALVLDSPEDVEKWKFTLTGGRPLTKQREKCGEPKCVFKNWLEIF 226  
 QY 408 EDDDKLEQIRKY---TSCAMLTGELKKALLIEVLOPLIAHQARRKEVTDEIVKEFMP 464  
 DB 227 EDDDK--KLKERYACKNELTCGCKRYLISKIQEFLKEHQRRRK-AEKLVKRRKYYG 283  
 QY 465 KLS 467  
 DB 284 KLA 286  
  
 RESULT 15  
 ID Q9HN66 PRELIMINARY; PRT: 380 AA.  
 AC Q9HN66;  
 DT 01-MAR-2001 (TREMBlrel. 16, Created)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE TRYPTOPHANYL-TRNA SYNTHETASE.  
 GN TRPS2 OR VNG2232G.  
 OS Halobacterium sp. (strain NRC-1).  
 OC Archaea; Euryarchaeota; Halobacteriales; Halobacteriaceae;  
 OC Halobacterium.  
 OX NCBI\_TaxID=64091;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20504483; PubMed=11016950;  
 RA Ng W.V., Kennedy S.P., Mahairas G.G., Bergquist B., Pan M.,  
 RA Shukla H.D., Lasky S.R., Baliga N.S., Thorsson V., Shrogha J.,  
 RA Swartzell S., Weir D., Hall J., Dahl T.A., Meli R., Goo Y.A.,  
 RA Leithausen B., Keller K., Cruz R., Danson M.J., Hough D.W.,  
 RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angelvine C.M., Dale H.,  
 RA Isenbarger T.A., Beck R.F., Pohlshroder M., Spudich J.L., Jung K.-H.,  
 RA Alam M., Frelas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,  
 RA Ebhardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.;  
 RT "Genome sequence of Halobacterium species NRC-1."  
 RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).  
 DR EMBL: AE005109; AAG20355.1; .  
 DR InterPro: IPR001994; Cytidylyltransf.  
 DR InterPro: IPR002305; tRNA-synt\_1b.  
 DR InterPro: IPR001412; tRNA-synt\_1.  
 DR InterPro: IPR002306; tRNA-synt\_trp.  
 DR Pfam: PF00579; tRNA-synt\_1b; 1.  
 DR PRINTS: PR01039; TRNASYNTHTRP.  
 DR PROSITE: PS00178; AA\_TRNA\_LIGASE\_1; UNKNOWN\_1.  
 KW Aminoacyl-tRNA synthetase; Complete proteome.

SQ SEQUENCE 380 AA; 41936 MW; 0F7B2B955386404F CRC64;  
  
 Query Match 21.0%; Score 535.5; DB 17; Length 380;  
 Best Local Similarity 35.2%; Pred. No. 1.8e-37;  
 Matches 142; Conservative 78; Mismatches 142; Indels 41; Gaps 17;  
  
 QY 80 ABEDEVDPWTVOTSSAKGIDYDKLLVIRGSSKIDELINRERATGQRPNHFRGIFFS 139  
 DB 3 ADGNQVTPYAVSSD---LDYEKLLARFGADELTDQARFP-----DHPLVNRGLPYA 53  
 QY 140 HDDMMQVLDAYENKKPFYLYTGRGSSSEAMVGHILPITKMLQDVENVPLVIOMTDE 199  
 DB 54 GRVDVDFPLTAGQS---IVTGVGPSG--PMHIGHAMVEFYFARRLQDEFGARYVPLSDE 108  
 QY 200 KYLMVDLTLDQAYGDAVE-NAKDIIACGFDIKFTLF---SDLDYM-GMSSGFYKNVYKI 254  
 DB 109 KYWFDQTPAEI-GDYLRANLBDLAVGFDPELTIIVYDTRADVLYPLATIFAGDV--- 164  
 QY 255 QKHVTFNVOYKGIFFGTDSDCIKISFPALQAPSPSNSFPQIFRDRTDIOCLIPCAIDOD 314  
 DB 165 -RHATLQNVYG-----EPDNGVGAFFYPAVQTAHL---LPQLVHG--EHETLVPIAVDOD 213  
 QY 315 PYFRMTRDVAPRIGYP--KPALHSTFEPALQAGQTKMSASDPNSSIFLDTAKQIKTKY 372  
 DB 214 PIVRVSBDVAAKARYPVGRGALLMQFLPSLAG-PGKMSSS-AGVYSIRLTPSPDVRK 271  
 QY 373 NKHAFSGGRDTEIEHRQFGNCDDVDSFMYLTFPLEDD-KLEQIRKYDTSGAMLTGELK 431  
 DB 272 KTHATTTGGRASVEEHRKAKGVPAEDVFPQYLSAFEPDDAELARLEREYRAGDLSGELK 331  
 QY 432 KALIEVLOPLIAHQARRKEVTDEIVKEFMPRKLSFPQOKLA 474  
 DB 332 DLAADRTFELAAHQRRRALGD--VTEALDAFRLTDERQRA 372

Search completed: October 24, 2002, 12:53:03  
 Job time : 31.2521 secs

us-09-813-718-10.rspt

```
QY 427 TGE LK KALIE-VLQPLAEHQARREVTDEIVKEEFTPKL 466
      ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 448 SGEVKDLAADCIIRFLVHDMERRKRTVTDVDVEECRINDI 488
```

### RESULT 13

ID	Q95295;	PRELIMINARY;	PRT;	136 AA.
AC	Q95295;			
DT	01-FEB-1997 (TrEMBLrel. 02, Created)			
DT	01-FEB-1997 (TrEMBLrel. 02, Last sequence update)			
DT	01-DEC-2001 (TrEMBLrel. 19, Last annotation update)			
DE	TRYPROPHANTYL-TRNA SYNTHASE (FRAGMENT).			
OS	Sus scrofa (Pig).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.			
OX	NCBI_TaxID=9823;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RP	TISSUE=SMALL INTESTINE;			
RC	Wintoroe A.K., Fredholm M., Davies W.;			
RA	"Evaluation and characterization of a porcine small intestine cDNA			
RT	library.";			

[illegible]

Db 61 PALGAGQTKMSXAPNSSIFLIDTAKQIKTKVKNHAFSGGRDIIIEHNKQFGGNCDDVVSF 120  
QY 401 MYLTFLEDDDKLEQI 416

Db 121 MYLTFLEDDDRXEQI 136

RESULT 14  
059584

AC 059584; (NCBI:taxid:1000000000) 0.7, Created)  
DT 01-AUG-1998 (TREMBLrel\_07, last sequence update)  
DT 01-AUG-1998 (TREMBLrel\_17, last annotation update)  
DT 01-JUN-2001 (TREMBLrel\_17, last annotation update)  
DE 301AA LONG HYPOTHETICAL TRYPTOPANYL-TRNA SYNTHETASE.  
GN PH1921.  
OS Pyrococcus horikoshii.  
OC Archaea: Euryarchaeota: Thermococcales: Thermococcaceae: Pyrococcus.  
OX NCBI\_TaxId=53953;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC  
RX MEDLINE=98344137; PubMed=9679194;  
RA Kawarabayashi Y., Sawada M., Horikawa H., Haikawa Y., Hino Y.,  
RA Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,  
RA Sakai M., Ogura K., Ohsuka R., Nakazawa H., Takamiya M., Ohnuku Y.,  
RA Funahashi T., Tanaka T., Kudoh Y., Yamazaki Y., Kushida N., Oguchi A.,  
RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,  
RA Masuchi K., Shizuya H., Kikuchi H.;  
RT "Complete sequence and gene organization of the genome of a hyper-thermophilic archaeobacterium, Pyrococcus horikoshii OT3.";

Best Local Similarity 64.9%; Pred. No. 1.1e-79;  
Matches 194; Conservative 39; Mismatches 64; Indels 2; Gaps 1;

OY 74 GPRATAEEDFVDPWTVQVSSAKG-IDYDKLIYRFGSSKIDKELINRIERATGQRPHHFLRGI 133  
DB 23 GCGVGEDEDEDRVTPWMTYTTKATGIDYDKLIYFGCRKIDDEELIAFERVYTHKASPMKR 82  
OY 134 RGIFFSHRDMNOVLDAVENKKRPFLYLTGRGSSSEAMHVGHLIPFTFKMLQDVFNPLVI 193  
DB 83 RGMFFAHRLTLALDRKEGKRPFLYLTGRGASSGSLHLHLVFIPTFKMLQGEVDPLVI 142  
OY 194 QMTDEKTYIMKDLTDQAVGDAVENAKDIIACGFDINKTFIESDLDVMGSSGFYKNVVK 253  
DB 143 QMTDDEKFLMKDKMDEAKMKARENKKDIISVGFDPKTFIENFDY--MCPEFYENIVK 200  
OY 254 IOKHVFNOVGKGFEGFTSDCIKGISFPALQAPSFNSFPQFERRTDIOCLIPCAIDPO 313  
DB 201 IMKVNTNARALFGFTPECCGKAAPFAVEAPCFASSFPQIFGRNDIPCLIPCAIDPO 260  
OY 314 DPFYMTRDVAPRIGYKPKPALHSTFPFALOGAQTMSASDPNSSIFLDTAKQIKTKV 372  
DB 261 DPFYMTRDVAPRILKASKSLIFSTFLPALTGATKMSASEPTCTIFLSDTAKQIKNKV 319

RESULT 10

OY976M1 PRELIMINARY; PRT; 406 AA.

AC O976M1; 01-DEC-2001 (TREMBlrel. 19, Created)  
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
DE PUTATIVE TRYPTOPHANYL-TRNA SYNTHETASE.  
GN ST0169.  
OS Sulfolobus tokodaii.  
OC Archaea; Crenarchaeota; Sulfolobales; Sulfolobaceae; Sulfolobus.  
OX NCBI\_TaxID=111955;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=JCM 10545 / 7;  
RX PubMed=11572479;  
RA Kawarabayashi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,  
RA Sekine M., Baba S.-I., Anai A., Kosugi H., Hosoyama A., Fukui S.,  
RA Nagai Y., Nishijima K., Otsuka K., Nakazawa H., Takamiya M., Kato Y.,  
RA Yoshizawa T., Tanaka T., Kudo Y., Yamazaki J., Kushida N., Oguchi A.,  
RA Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A.,  
RA Oshima T., Kikuchi H.;  
RT "Complete genome sequence of an aerobic thermophilic  
RT Crenarchaeon, Sulfolobus tokodaii strain7.";  
RL DNA Res. 8:123-140(2001).  
DR EMBL; AP000981; BAB65126.1; -;  
KW Aminoacyl-tRNA synthetase; Hypothetical protein; Complete proteome.  
SQ SEQUENCE 406 AA; 47979 MW; 276CDB506B94B3 CRC64;

Query Match 35.7%; Score 910.5; DB 17; Length 406;

Best Local Similarity 50.5%; Pred. No. 1.5e-69;  
Matches 189; Conservative 63; Mismatches 99; Indels 23; Gaps 9;

OY 82 EDF-VDPWTVQVSSAKG-IDYDKLIYRFGSSKIDKELINRIERATGQRPHHFLRGI 139  
DB 28 GGFNTVPMVEV-----KGKDYDKLIYQFGTKITSELKIKSIINDELHVMKRNVFS 82  
OY 140 HRDMNOVLDAVENKKRPFLYLTGRGSSSEAMHVGHLIPFTFKMLQDVFNPLVIQMTDE 199  
DB 83 HRDLIDLVLADYDQKGFPLYTGRAPSL-GMHIGHLIPFTFKMLQDKFNVNLIYETDDE 141  
OY 200 KTLWK-DLTDQAVGDAVENAKDIIACGFDINKTFIESDLDVMGSSGFYKNVVKQ 258  
DB 142 KTRMNEDEYTLDTQROMAYNIDIIAVGFNPDKTFLEQTEYI--RNMYPIAKIATKL 198  
OY 259 TPNQVKGIFGFTSDCIKGISFPALQAPSFNSFPQIFGRNDIPCLIPCAIDODPYR 318  
DB 199 TSEVATATGDLTSSNIGIITWPAIDIAFT-----MEKR---RCLIPAGIDODPYR 248

OY 319 MTRDVAAPRIGYKPKPALHSTFPFALOGAQTMSASDPNSSIFLDTAKQIKTKVKNHAPS 378  
DB 249 LQRDIENSLGYKAAQIHISKFLPLTLGPECKSSSQPETAIVLTDDPKVVERIKMYARS 308  
OY 379 GGRDTIEHRQGNCDVVSFMYLTFLE-DDEKIQIRKQYTSAMTGLKALIV 437  
DB 309 GGPRTIELHRKYGNDIDVVSQWLYMFEPDPDNKIKKIEEDYRSGALLTGELKOLIEK 368  
OY 438 LQPLAEHQARKE 451  
DB 369 LNDLEEHROKRE 382

RESULT 11

OY972X0 PRELIMINARY; PRT; 386 AA.

AC O972X0; 01-OCT-2001 (TREMBlrel. 18, Created)  
DT 01-OCT-2001 (TREMBlrel. 18, Last sequence update)  
DT 01-OCT-2001 (TREMBlrel. 18, Last annotation update)  
DE TRYPTOPHANYL-TRNA SYNTHETASE (TRPS) (EC 6.1.1.2).  
GN TRPS.  
OS Sulfolobus solfataricus.  
OC Archaea; Crenarchaeota; Sulfolobales; Sulfolobaceae; Sulfolobus.  
OX NCBI\_TaxID=2287;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ATCC 35092 / DSM 1617 / P2;  
RX MEDLINE=11427726; PubMed=11427726;  
RA She O., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,  
RA Aweyer M.J., Chan-Welher C.C.-Y., Clausen I.G., Curtis B.A.,  
RA De Moers A., Erasus G., Fletcher C., Gordon P.M.K.,  
RA Heikamp-de Jong I., Jeffries A.C., Kozeira C.J., Medina N., Peng X.,  
RA Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,  
RA Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,  
RA Garrett R.A., Ragan M.A., Sensen C.W., Van der Oost J.;  
RT "The complete genome of the crenarchaeon Sulfolobus solfataricus P2.";  
RL Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).  
DR EMBL; AE006677; AAK40778.1; -;  
DR InterPro: IPR002305; trna-synt\_1b.  
DR InterPro: IPR002306; trna-synt\_trp.  
DR Pfam: PF00579; trna-synt\_1b; 1.  
DR PRINTS: PRO1039; TRNASYNTHTRP.  
KW Aminoacyl-tRNA synthetase; Ligase; Complete proteome.  
SQ SEQUENCE 386 AA; 45448 MW; 3F4DA8C9B8D219DF CRC64;

Query Match 35.6%; Score 907; DB 17; Length 386;

Best Local Similarity 48.6%; Pred. No. 2.8e-69;  
Matches 191; Conservative 68; Mismatches 110; Indels 24; Gaps 10;

OY 78 TEAEDEFPWTVQVSSAKG-IDYDKLIYRFGSSKIDKELINRIERATGQRPHHFLRGI 136  
DB 6 TMRDEFTVPMVEV-----KGKDYDKLIYQFGTKITSELKORINLAGDL-HVMKRNV 59  
OY 137 FFSHRDMNOVLDAVENKKRPFLYLTGRGSSSEAMHVGHLIPFTFKMLQDVFNPLVIQMT 196  
DB 60 FFSHRDIDLVDYDEKSKGFPLYTGRAPSL-GMHIGHLIPFTFKMLQDKFNANLIYET 118  
OY 197 DDEKTYIMK-DLTDQAVGDAVENAKDIIACGFDINKTFIESDLDVMGSSGFYKNVVKQ 255  
DB 119 DDEKTYIMNEFTLDTQRSNAYNIDIIAVGFNPDKTFLEQTEYI--RNMYPIATVKA 175  
OY 256 KHTFNOVKGIFGFTSDCIKGISFPALQAPSFNSFPQIFGRNDIPCLIPCAIDODPYR 315  
DB 176 KLTSEVATATGDLTSSNIGIITWPAIDIAFT-----MEKR---RCLIPAGIDODP 225  
OY 316 YFRMTDVAAPRIGYKPKPALHSTFPFALOGAQTMSASDPNSSIFLDTAKQIKTKVKNH 375  
DB 226 YWRLORDIAESLGYYKAAQIHISKFLPLTLGPECKSSSQPETAIVLTDDPKVVERIKMY 285  
OY 376 AFSGGDTIEHRQGNCDVVSFMYLTFLEDDD-KEDQIRKQYTSAMTGLKALIV 434  
DB 286 AFSGGQPTIELHRKYGNDIDVVSQWLYYFPEEDDNRIKIEEYRSGKMLTGLKOLTL 345

"The genome sequence of Drosophila melanogaster.";

RL Science 287:2185-2195(2000).  
 DR EMBL: AE003682; AAC22136.1; .  
 DR FlyBase: FBgn0010803; Aats-trp.  
 DR InterPro: IPR002305; tRNA-synt\_1b.  
 DR InterPro: IPR001412; tRNA-synt\_1.  
 DR InterPro: IPR002306; tRNA-synt-trp.  
 DR Pfam: PF00579; tRNA-synt\_1b; 1.  
 DR PRINTS: PR01039; TRNASYNTHTRP.  
 DR PROSITE: PS00178; AA\_TRNA\_LIGASE\_1; 1.  
 DR SEQUENCE 430 AA; 47971 MW; 2938EECC69E979F CRC64;

Query Match 54.1%; Score 1378; DB 5; Length 430;  
 Best Local Similarity 61.4%; Pred. No. 1.5e-109;

Matches 262; Conservative 64; Mismatches 85; Indels 16; Gaps 3;

OY 57 DYKADCPGPNPAPTSNHGPD-----TEAE-----EDFVDPWTVQTSSAKGIDYD 101  
 DB 3 DKEKEVVEGVETLNGKRDAPVEGTDAQOEGATAPTEVDVDPWNVASSNDAGVDYD 62  
 OY 102 KLIVRSGSKIDKELINRERATGPRPHHFLRGIFESHROMNQVLDAYENKKPPYLTYG 161  
 DB 63 KLIRKFGSSKIDELIAREKITGKPAHHFIRGMPFSHRDLHTLTTLREGKPPYLTYG 122  
 OY 162 RGPSEAMHGHLPFIETKWLQDVFNPLVQMTDDEKYLKDLTLDOAYGDAVENAKD 221  
 DB 123 RGPSSGSLHGVHLVPIFMTKWLOETFDVPLVQLTLDDEKTLMKDLKVEDAIKLGRENKD 182  
 OY 222 ITACGPDINKTIFESDLDYMGSSGFEYKNVVKIQKHVTFNQVKGIFGFTDSCIGKISFP 281  
 DB 183 IVAIFEDVKTFFIFNNLEFEVKGCPAMYONIIRIQKCTFENQVKGIFGFTDSCILKIGRP 242  
 OY 282 ALOAPSPNSPPOIFRQDITDQCLIPCAIDDPFRMRDVAAPRGVAPKALLHSTFP 341  
 DB 243 AAOAAPALSTSTPFLFGN-KVHCLIPCAIDDPFRMRDVAAPRGVAPKALLHSTFP 301  
 OY 342 ALOGATKMSASDPNSSIFLDTAKQIKTKVKNHAFSGGRDTEEHROFGMCDVDSFM 401  
 DB 302 ALOGAKTKMSADQNSAVVLTDTPKQIKKINKYAFSGSRVSEEHKRGKGVPEVDVSYQ 361  
 OY 402 VLTFFLEDDEKLEQIRKDYTSGLMTGELKALILEVLOPLAHOARRKEVTDEIVKEFM 461  
 DB 362 LKFFLEDDEKLEEVAVVSKGEMLTGELIKLAVETLPIVBOHQAARKLITDEVLDKYF 421  
 OY 462 TPRKLSF 468  
 DB 422 ELRPLKF 428  
 RESULT 8  
 O9SR15 PRELIMINARY: PRT; 402 AA.  
 AC O9SR15  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE PUNATIVE TRYPTOPHANYL-TRNA SYNTHETASE.  
 GN F7018.7.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eucosids II; Brassicales; Brassicaceae; Arabidopsids.  
 OC NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. COLUMBIA;  
 RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B.,  
 RA Ronning C.M., Koo H., Fujii C.Y., Uteback T.R., Barstead M.E.,  
 RA Bowman C.L., White O., Nierman W.C., Fraser C.M.;  
 RT "Arabidopsis thaliana chromosome IIT BAC F7018 genomic sequence.";  
 RL Submitted (JAN-2001) to the EMBL/Genbank/DBJ databases.  
 DR EMBL: AC011437; AAF04890.1; .  
 DR InterPro: IPR002305; tRNA-synt\_1b.

DR InterPro: IPR001412; tRNA-synt\_1.  
 DR InterPro: IPR002306; tRNA-synt-trp.  
 DR Pfam: PF00579; tRNA-synt\_1b; 1.  
 DR PRINTS: PR01039; TRNASYNTHTRP.  
 DR PROSITE: PS00178; AA\_TRNA\_LIGASE\_1; 1.  
 KW Aminoacyl-tRNA synthetase.  
 SQ SEQUENCE 402 AA; 45754 MW; B9BEA75EE56CD15 CRC64;

Query Match 51.1%; Score 1301; DB 10; Length 402;  
 Best Local Similarity 60.3%; Pred. No. 5.3e-103;  
 Matches 240; Conservative 66; Mismatches 88; Indels 4; Gaps 2;

OY 76 DATEAE--EDFVDPWTVQTSSAKGIDYDKLIYRFGSSKIDKELINRERATGPRPHHFLR 133  
 DB 7 DERAEASSQVYVNPWEVSAKDGKIDYDKLIDKFCGQRDESLDRVQLRTSRQGHVFLR 66  
 OY 134 RGIFFSHRDMNOVLDAYENKKRPYLYTGRGPSEAMHGHLPFIETKWLQDVFNPLVI 193  
 DB 67 RSVFPAHRDFNELLDAVERGDFLYLTGRGPSSSEALHGHLPFMFTYXLOEAFKVPPLVI 126  
 OY 194 QMTDDEKYLKDLTLDOAYGDAVENAKDILACGPDINKTIFESDLDYMGSSGFEYKNVVK 253  
 DB 127 QLTDEKSLMKNLVSESORLARENKADILACGPDVTKTIFESDPIYG--GAFYKNVVK 184  
 OY 254 IOKHVTFNQVKGIFGFTDSCIGKISFPALQAPSPNSPPOIFRDRDITDQCLIPCAIDQ 313  
 DB 185 VCKCVTLKAMGIFGSGEDPIAKLSFPVQAVPSPSPFPLPGKDWLRCLIPCAIDQ 244  
 OY 314 DPFPMTRDVAAPRGVAPKALLHSTFPALQAGQTKMSASDPNSSIFLDTAKQIKTKVN 373  
 DB 245 DPFPMTRDVAAPRGVAPKALLHSTFPALQAGENKMSASDPNSAIYVYDSAKDKKNIN 304  
 OY 374 KHAFFSGGRDTEEHROFGMCDVDSFMVLTFFLEDDEKLEIRKDYTSGLMTGELKKA 433  
 DB 305 RAFFSGGQDSTIKHRELGNLEVDLPVKYLSFFLEDDEKLEIKETGEGRMALTEGVKKR 364  
 OY 434 LIEVLOPLAHOARRKEVTDEIVKEFMTPRKLSFDFQ 471  
 DB 365 LFEVLTEIVERRRARRAANTDEMVAFAFNAVRLPMPFE 402

RESULT 9  
 O9UIR2 PRELIMINARY: PRT; 324 AA.  
 AC O9UIR2  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 DE Y80D3A.1 PROTEIN.  
 GN Y80D3A.1.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OC NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Harris B.R.;  
 RT Submitted (SEP-1999) to the EMBL/Genbank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99069613; PubMed=9851916;  
 RA none;  
 RT "Genome sequence of the nematode C. elegans: A platform for  
 RT investigating biology.";  
 RL Science 282:2012-2018(1998).  
 DR EMBL: AL132853; CAB60439.1; .  
 DR InterPro: IPR002305; tRNA-synt\_1b.  
 DR InterPro: IPR002306; tRNA-synt-trp.  
 DR Pfam: PF00579; tRNA-synt\_1b; 1.  
 DR PRINTS: PR01039; TRNASYNTHTRP.  
 DR SEQUENCE 324 AA; 36289 MW; 66687E6E420ECC12 CRC64;

Query Match 40.3%; Score 1027; DB 5; Length 324;

OY	162	RBPSSAMHVGHLIPFTITKMLQDVFNPLQMTDDEKYLMLKQTLTDOAVGDAVENAKD	223
Db	123	RBPSSGSLHVGHLVFFITKMLQLETFVPLVLTQLTDEKTLMLKQEDAKIKGRENAKD	122
OY	222	IIACGFIDINKTFFESDLDYMGMSGCFYKNVYIKQHVTFMNVKGIFFGFTSDICIGKISFP	281
Db	183	IYAIIEFVNKTEFFIEFNLEEFVCKEPCAMYQNIIRIQCKVTFMNVKGIFFGFSGDITGIGFP	242
OY	282	AIQAPSPNSNPQEFRRRTDIOCLIPADIDDDYFRRTRVAPRIGYKPRALLHSTFFP	341
Db	243	AAQAPALISTFPEFTEGNR-KVCHCLIPADIDDDYFRRTRVAPRIGYKPRALLHSTFFP	301
OY	342	ALQGAOTMSASDPPSSIFLDTFAKQIKTKYNKAHAFSGGRDTEIEHSGPCNDVDVSFM	401
Db	302	ALQGAKTMSASDQNSAYVLTDFRKQIKNKAKAFSGGRGTVEHKKLGSPVENDVSQ	361
OY	402	YLTFFLEDDDKLEQIRKQYTSGCAMLTGELKALIEVLQPLIAEQARRKEVDEIYEFM	461
Db	362	LKLFLEDDAKLEEVRAVYSGEMLTGEIKLAEVETLPTVECHOARKLITTEVDLKYF	421
OY	462	TPRKLSF	468
Db	422	ELRPLKF	428
RESULT 6			
Q9U4Y0			
ID	Q9U4Y0	PRELIMINARY;	PRT; 420 AA.
AC	Q9U4Y0.		
DT	01-MAY-2000 (TREMBLrel.13, Created)		
DT	01-MAY-2000 (TREMBLrel.13, Last sequence update)		
DT	01-JUN-2001 (TREMBLrel.17, Last annotation update)		
DE	TRYPTOPHANYL-TRNA SYNTHETASE (FRAGMENT).		
GN	AATS-TRP OR CG9735.		
OS	Drosophila melanogaster (Fruit fly).		
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;		
OC	Eukaryota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;		
OC	Ephydroidea; Drosophilidae; Drosophila.		
OX	NCBI_TaxID=7227;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=99250164; PubMed=10233165;		
RA	Seshalain P., Andrew D.J.;		
RT	"WMS-85D: A tryptophanyl-tRNA synthetase expressed to high levels in		
RT	the developing Drosophila salivary gland."		
RL	Mol. Biol. Cell 10:1595-1608(1999).		
DR	EMBL: AF125157; AAF20167.1;--		
DR	FLYBase; FBgn0010803; Aats-trp.		
DR	InterPro; IPR002305; tRNA-synt_1b.		
DR	InterPro; IPR001412; tRNA-synt_1.		
DR	InterPro; IPR002306; tRNA-synt_trp.		
DR	Pfam; PF00579; tRNA-synt_1b; 1.		
DR	PRINTS; PR01039; TRNASYNTHTRP.		
DR	PROSITE; PS00178; AA_TRNA_LIGASE_I; 1.		
KW	Aminoacyl-tRNA synthetase.		
FT	NON_TER		
FT	1		
SO	SEQUENCE		
	420 AA; 46975 MW; 4A4F708BAFA426AEB6 CRC64;		
Query Match	54.28; Score 1379.5; DB 5; Length 420;		
Best Local Similarity	62.78; Pred. No.1.le-109;		
Matches 257; Conservative	62; Mismatches 90; Indels 1; Gaps		
OY	59	KADCPGNPAFTSNHGPAETAEEDFVDPWVYFSSAKGIDYKLIYRFSSKIDKELIN	118
Db	10	KPDAPVETGIDAQAQEGCATAPTEDVDVDPWVVASNAGVDYKLIYRFSSKIDDELLA	69
OY	119	KIERATGCRPHHFLRGIFFSHRDMNOYLDAENKKPFYLYTGCGSSSEAMHVGHLIPT	178
Db	70	RFEKITGCRPHHFLIRGFFSHRDLHLLTLEQGRPFYLYTGCGSSGSLHVGHLVPTI	129
OY	179	FTKMLQDVFNPLVLTQMTDDEKYLMLKQTLTDOAVGDAVENAKDIIACGFIINTFFESDL	238

Db	130	MTKMLQEFDFVPLVIGTLDDEKTLMKRLAKVEDAIKLGRENAKOIVAGIDVAKTFLFNNL	188
Qy	239	DYMGSSGCFKRVYKIOKHTFNQKCIIEFTDSDCIKISPAIOAAPSFSNPQIFR	298
Db	190	EFVKGCPAMVONIRIORKCVTFNVOKSIFGFGSDIILGKGFPAOAAVAISTPEPFIIG	249
Qy	299	DRTDIQICIPCAIDODDYPFRMTNRVAPRIGYPRKPAIIHSFPFPAILOGAOMKMSASPNSS	358
Db	250	NR-KVHCLIPCAIDODDYPFRMTNRVAPRIGYPRKPAIIHSFPFPAILOGAOMKMSASPNSS	308
Qy	419	DYTSGAMITGELKALIEVLQPLAEHQARKEVTEIYAKEMTPKLSF	468
Db	369	AYSQGEMLTGEIKRLAVERITPEIYEQHQAARKLITPELDVKYELPLKFX	418
RESULT 7			
QyVHG2	09VHG2	PRELIMINARY;	PRT; 430 AA.
AC	Q9VHG2;		
DT	01-MAY-2000 (TREMBlrel. 13, Created)		
DT	01-MAY-2000 (TREMBlrel. 13, last sequence update)		
DT	01-JUN-2001 (TREMBlrel. 17, last annotation update)		
DE	AATS-TRP PROTEIN.		
GN	AATS-TRP OR CG9735.		
OS	Drosophila melanogaster (Fruit fly).		
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;		
CC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;		
CC	Ephydroidea; Drosophilidae; Drosophila.		
OX	NCBI_taxid=7227;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-BERKELEY.		
FX	MEDLINE=20196006; PubMed=10731132;		
RA	Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,		
RA	Amaratilake P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,		
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,		
RA	Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,		
RA	Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,		
RA	Wan K.H., Doyle C., Baxter E.G., Holt G., Nelson C.R., Miklos G.L.G.,		
RA	Abtill J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,		
RA	Ballew R.W., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,		
RA	Beeoon K.Y., Benos P.V., Bertman B.P., Bhandari D., Bolshakov S.,		
RA	Botkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,		
RA	Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,		
RA	Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,		
RA	de Pallos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,		
RA	Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,		
RA	Fowler K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischman W.,		
RA	Poser C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K.,		
RA	Glodak A., Gong F., Gottrell J.H., Gu Z., Guan P., Harris M.,		
RA	Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,		
RA	Hoskin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,		
RA	Jatallil M., Kalush F., Kapen G.H., Ke Z., Kennison J.A., Ketchum K.A.,		
RA	Kimmel B.E., Kohira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,		
RA	Liao X., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,		
RA	Lian X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,		
RA	Merkulov G., Mishina N.V., Moberly C., Morris J., Moshrefi A.,		
RA	Mount S.M., Moy M., Murphy B., Murphy L., Munzy D.M., Nelson D.L.,		
RA	Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,		
RA	Palazzo M., Pittman G.S., Pan S., Pollard J., Puri V., Reen M.G.,		
RA	Ralnet K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,		
RA	Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,		
RA	Spieler E., Spreading A.C., Stapleton M., Strong R., Sun E.,		
RA	Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,		
RA	Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,		
RA	Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,		
RA	Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,		
RA	Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,		
RA	Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;		

Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
RA Suuki H., Toyooka K., Wang K.H., Weltz C., Whitaker C., Wilming L.,  
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohlsuki S.,  
RA Hayashizaki Y.;  
RT "Functional annotation of a full-length mouse cDNA collection.";  
RL Nature 409:685-690(2001).  
DR EMBL: AK007754; BAB25235.1; -  
DR MGI: 104630; Mavs.  
DR InterPro: IPR002305; tRNA-synt\_1b.  
DR InterPro: IPR001412; tRNA-synt\_1.  
DR InterPro: IPR002306; tRNA-synt\_trp.  
DR Pfam: PF00579; tRNA-synt\_1b; 1.  
DR PRINTS: PR01039; TRNASYNTHTRP.  
DR PROSITE: PS00178; AA\_TRNA\_LIGASE\_1; 1.  
SO SEQUENCE 329 AA; 37613 MW; 701E702DC244CA42 CRC64;

Query Match 64.0%; Score 1629; DB 11; Length 329;  
Best Local Similarity 92.1%; Pred. No. 3,2e-131;  
Matches 303; Conservative 17; Mismatches 9; Indels 0; Gaps 0;

QY 143 MOVDAVENKKPFYLYTGRGSSSEAMHVLIPFTKMLQDVNPLVLIQMTDDEKYL 202  
DB 1 MNOUIDAENKKPFYLYTGRGSSSEAMHVLIPFTKMLQDVNPLVLIQMTDDEKYL 60  
QY 203 WKDLTLDAQYGDVAENAKDIIACGFDINKFTFSDLDYMGSSGFYKNVVKIQKHVTENQ 262  
DB 61 WKDLTLDAQYGVENAKDIIACGFDINKFTFSDLEVMGSGPYRNVVKIQKHVTENQ 120  
QY 263 VKGIFGFDSQICIKISFPALQAAFSFNSPFOIFRDRTDIOCLIPCAIDDPYRMRD 322  
DB 121 VKGIFGFDSQICIKISFPALQAAFSFNSPFIKFRDRTDIOCLIPCAIDDPYRMRD 180  
QY 323 VAPRIGYKPKPALHSTFFPALQAGTQKMSASDPNSSIFLTDTAKOIKRKVNKHAESGSD 382  
DB 181 VAPRIGHKPKPALHSTFFPALQAGTQKMSASDPNSSIFLTDTAKOIKRKVNKHAESGSD 240  
QY 383 TIEEHROFGNCDVDSFMYLTFLEDDDKLEQIRKDYTSGLMTELKALIEVLOPLI 442  
DB 241 TIEEHROFGNCEVDVSFMYLTFLEDDDKLEQIRKDYTSGLMTELKALIEVLOPLI 300  
QY 443 AEHQARRKEVTDEYKEFMTPRKISFDFQ 471  
DB 301 AEHQARRKAVTEYKEFMTPRKISFDFQ 329

RESULT 4  
070184 PRELIMINARY; PRT; 305 AA.  
ID 070184; PRT; 305 AA.  
AC 070184;  
DT 01-AUG-1998 (TREMBLrel. 07, Created)  
DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE TRYPTOPHAN-TRNA SYNTHETASE (FRAGMENT).  
OS Cavia porcellus (Guinea pig).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Hystriocognathia; Cavidae; Cavia.  
OX NCBI\_Taxid=10141;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-HARTLEY; TISSUE-SPLEEN;  
RA Yang D., Goto R., Watanabe N., Kobayashi Y.;  
RT "Identification and Cloning of Genes Whose Expressions are Elevated  
RT during DMCB-Induced Guinea Pig Skin Delayed-type Hypersensitivity  
RT Reaction.";  
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AB012222; BAA25288.1; -  
DR InterPro: IPR002305; tRNA-synt\_1b.  
DR InterPro: IPR002306; tRNA-synt\_trp.  
DR Pfam: PF00579; tRNA-synt\_1b; 1.  
DR PRINTS: PR01039; TRNASYNTHTRP.  
KW Aminoacyl-tRNA synthetase.  
FT NON\_TER 1  
SQ SEQUENCE 305 AA; 34838 MW; 0937164333780EB7 CRC64;

Query Match 60.3%; Score 1537; DB 11; Length 305;  
Best Local Similarity 94.4%; Pred. No. 2,2e-123;  
Matches 288; Conservative 11; Mismatches 6; Indels 0; Gaps 0;

QY 167 EAMHGHILIPFTKMLQDVNPLVLIQMTDDEKYLKMDLTLDAQYGDVAENAKDIIACG 226  
DB 1 EAMHGHILIPFTKMLQDVNPLVLIQMTDDEKYLKMDLTLDAQYGVLENAKDIIACG 60  
QY 227 FDINKTFIFSDLDYMGSSGFYKNVVKIQKHVTENQVGIIGFTDSDIGKISFPAIOAA 286  
DB 61 FDINKTFIFSDLEVMGSSGFYKNVVKIQKHVTENQVGIIGFTDSDIGKISFPAIOAA 120  
QY 287 PSFNSPFOIFRDRTDIOCLIPCAIDDPYRMRDVAAPRIGYKPKPALHSTFFPALOGA 346  
DB 121 PSFNSPFOIFRDRTDIOCLIPCAIDDPYRMRDVAAPRIGYKPKPALHSTFFPALOGA 180  
QY 347 QTKMSASDPNSSIFLTDTAKOIKRKVNKHAESGSDTIEEHROFGNCDVDSFMYLTF 406  
DB 181 QTKMSASDPNSSIFLTDSAKOIKRKVNKHAESGSDTIEEHROFGNCDVDSFMYLTF 240  
QY 407 LEDDDKLEQIRKDYTSGLMTELKALIEVLOPLIAHQARRKEVTDEYKEFMTPRKL 466  
DB 241 LEDDDKLEQIRKDYTSGLMTELKALIEVLOPLIAHQARRKEVTDEYKEFMTPRKL 300  
QY 467 SFDFQ 471  
DB 301 SFDFQ 305

RESULT 5  
0904Y1 PRELIMINARY; PRT; 430 AA.  
ID 0904Y1;  
AC 0904Y1;  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE TRYPTOPHANYL-TRNA SYNTHETASE.  
GN AATS-TRP OR CG9735.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_Taxid=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99250164; PubMed=10233165;  
RA Seshalah P., Andrew D.J.;  
RT "WRS-85D: A tryptophanyl-tRNA synthetase expressed to high levels in  
RT the developing Drosophila salivary gland.";  
RL Mol. Biol. Cell 10:1595-1608(1999).  
DR EMBL: AF125156; AAF20166.1; -  
DR FLYBase: FBgn010803; Aats-trp.  
DR InterPro: IPR002305; tRNA-synt\_1b.  
DR InterPro: IPR001412; tRNA-synt\_1.  
DR InterPro: IPR002306; tRNA-synt\_trp.  
DR Pfam: PF00579; tRNA-synt\_1b; 1.  
DR PRINTS: PR01039; TRNASYNTHTRP.  
DR PROSITE: PS00178; AA\_TRNA\_LIGASE\_1; 1.  
KW Aminoacyl-tRNA synthetase.  
SO SEQUENCE 430 AA; 47985 MW; 2E3F9E9BC1E9979A CRC64;

Query Match 54.3%; Score 1382; DB 5; Length 430;  
Best Local Similarity 61.6%; Pred. No. 6,8e-110;  
Matches 263; Conservative 63; Mismatches 85; Indels 16; Gaps 3;

QY 57 DYKADCPGNPAPTSNMGPD-----TEAE-----EDFYDPMVTVQTSAGKIDYD 101  
DB 3 DTKETVVGCVGALTLNGPRDAEPVETGDAQAQBEGATPFDYDVPWVAVSSNDAGVYD 62  
QY 102 KLIYFGSSKIDKELINIERATGQRPHPLRGLGIFESHROMNOVLDAVENKKRPVLYTG 161  
DB 63 KLIKFGSSKIDDELIAFEKITGKPAHHFTRGWFESHRLHTILTLREGGKRPVLYTG 122



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 24, 2002, 12:49:45 ; Search time 26.2521 Seconds  
(without alignments)  
3189.442 Million cell updates/sec

Title: US-09-813-718-10

Perfect score: 2547  
Sequence: 1 MPNSEPASLLEFNSIATG.....KLSFDQKLAALHHHHHH 484

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

SPTREMBL\_19:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mmc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2227.5	87.5	475	11	Q9DC65	Q9dc65 mus musculus
2	2223.5	87.3	481	11	Q9J558	Q9j558 mus musculus
3	1629	64.0	329	11	Q9D8R9	Q9d8r9 mus musculus
4	1537	60.3	305	11	070184	070184 cavia porce
5	1382	54.3	430	5	Q9U4Y1	Q9u4y1 drosophila
6	1379.5	54.2	420	5	Q9U4Y0	Q9u4y0 drosophila
7	1378	54.1	430	5	Q9VHG2	Q9vhg2 drosophila
8	1301	51.1	402	10	Q9SR15	Q9sr15 arabidopsis
9	1027	40.3	324	5	Q9U1R2	Q9u1r2 caenorhabdi
10	910.5	35.7	406	17	Q976M1	Q976m1 sulfolobus
11	907	35.6	386	17	Q97ZX0	Q97zx0 sulfolobus
12	800.5	31.4	490	5	Q9U1F5	Q9u1f5 leishmania
13	682	26.8	136	6	Q9S295	Q9s295 sus scrofa
14	626.5	24.6	301	17	Q9S584	Q9s584 pyrococcus
15	535.5	21.0	380	17	Q9HN66	Q9hn66 halobacteri
16	428	16.8	136	6	Q9T588	Q9t588 bos taurus

17	310	12.2	111	5	Q95YL8	Q95yl8 encephalito
18	299.5	11.8	426	17	Q978Y8	Q978y8 thermoplasma
19	294	11.5	157	5	Q9U533	Q9u533 trypanosoma
20	290	11.4	109	4	Q9UD15	Q9ud15 homo sapien
21	274.5	10.8	513	17	Q9HN83	Q9hn83 halobacteri
22	268.5	10.5	426	17	Q9H1W5	Q9h1w5 thermoplasma
23	188	7.4	364	17	Q9YA64	Q9ya64 aeropyrum p
24	185.5	7.3	341	16	Q97N42	Q97n42 streptococc
25	183	7.2	340	16	Q99XH4	Q99xh4 streptococc
26	181.5	7.1	331	17	Q97921	Q97921 thermoplasma
27	174.5	6.9	321	16	Q9CJDI	Q9cjd1 lactococcus
28	168.5	6.6	895	10	Q9SGN2	Q9sgn2 arabidopsis
29	167	6.6	351	16	Q9RVD6	Q9rvd6 deinococcus
30	163	6.4	460	10	P93018	P93018 arabidopsis
31	159	6.2	33	4	P78534	P78534 homo sapien
32	158.5	6.2	375	17	Q9V027	Q9v027 pyrococcus
33	155	6.1	1714	5	Q95TJ3	Q95tj3 drosophila
34	154.5	6.1	102	1	Q07119	Q07119 halobacteri
35	152.5	6.0	682	5	Q9N9B8	Q9n9b8 leishmania
36	149.5	5.9	327	17	Q9HN62	Q9hn62 halobacteri
37	149.5	5.9	408	10	P93363	P93363 nicotiana t
38	147	5.8	528	11	Q91WQ3	Q91wq3 mus musculus
39	146	5.7	375	17	Q58739	Q58739 pyrococcus
40	145	5.7	525	5	Q9VW60	Q9vw60 thermoplasma
41	144	5.7	332	17	Q9HKT3	Q9hkt3 drosophila
42	140.5	5.5	1149	5	Q23315	Q23315 caenorhabdi
43	139	5.5	294	17	Q96YV3	Q96yv3 sulfolobus
44	138.5	5.4	449	5	Q9NG36	Q9ng36 drosophila
45	138.5	5.4	449	5	Q9NG35	Q9ng35 drosophila

## ALIGNMENTS

RESULT 1  
ID Q9DC65 PRELIMINARY; PRT; 475 AA.  
AC Q9DC65;  
DT 01-JUN-2001 (TREMBlrel. 17, Created)  
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)  
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
DE ADULT MALE LONG CDNA, RIKEN FULL-LENGTH ENRICHED LIBRARY,  
DE CLONE:1200002C07, FULL INSERT SEQUENCE.  
GN WARS.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=LUNG;  
RX MEDLINE=21085660; PubMed=11217851;  
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,  
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiya H.,  
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
RA Schirral L.M., Staabli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsi G.,  
RA Blake J., Boffelli D., Bojunga N., Carlini P., de Bonaldo M.F.,  
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mommaerts P.,  
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilmink L.,  
RA Wysshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,  
RA Hayashizaki Y.,  
RT "Functional annotation of a full-length mouse cDNA collection.",  
RL Nature 409:685-690(2001).  
DR EMBL; AK004541; BAB23357.1; -.





Query Match	6.8%;	Score 172;	DB 1;	Length 394;
Best Local Similarity	22.7%;	Pred. No. 2.1e-06;		
Matches	88;	Conservative	62;	Mismatches 132;
			Indels	106;
			Gaps	22;

```

QY 141 RDMNOVL-----DAYE-NKKPFYLYGRGSSAMHVGHJIPFIKTYLODFNV--PL 131
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 18 KNIDVNLNPQIKYIDVEYQKRHLKLYGTAPTGRH-HCYFVH--MTR-LADFLKAGCEV 73
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 192 VIOMTDEKYL-----WKDITLD--QAYDAVENADIIACGFIDINKTF 233
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 74 TVLLADLHAFLDNMKAPLEVYVNRKAYTELTIKALISINWPIEKLKFVVGSSYQTLPOV 133
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 234 IFSDDIYMGSSGGYKAVNK-----IKNHTFNQVKIFGFTSDCIKISPAIQAPS 288
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 134 ---TMDIFRLSINIYSONDAKRAQADVYKQVAPLISGLI-----YELMOA--- 175
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 289 FSNSEPOIFNRRTDIOCLIPALIDQDYPFMRTRVAVARIGYPRKALLSHFFPAL--QGAQ 347
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 176 -----LDEQFLDVDCQFC-GVDQKIFVLAENHNLISLGYKRAHLNMPVGLQAGG- 226
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 348 TRMSASDPNSSIFLTPAKOIKTYVNKHAHSGGHDITEH-----ROFGG 392
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 227 -KMSASDPNKSITDLLEERPKQKKKINAFCSPG--ANEENGILSFVQYVYAPLOELKFGI 283
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 393 NCDVDVSEMYLTFLELDDK-----LEQIKDYTSCAMLTGEIK---KALIEVIO 439
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 284 N-----HEEFDIPDEKFGGRTYKSEEMKLAKEEKLSPDPDIKIGVADAINELLE 335
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 440 PLIAHQARKEVTDLIVKEH--MTPRK 465
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 336 P-IRQEFANKEFOEASEKEGVPATPQK 362
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

Search completed: October 24, 2002, 12:51:42  
Job time : 9.87395 secs

AC Q46127; 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA ligase)  
DE (TrpRS).  
GN TRPS OR TRSA.  
OS Clostridium longisporum.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;  
OC Clostridium.  
OX NCBI\_TaxID=1533;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=B6405;  
RA Brown G.D., Thomson J.A.;  
RL Submitted (DEC-1995) to the EMBL/Genbank/DBJ databases.  
CC -1- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +  
CC diphosphate + L-tryptophanyl-tRNA(Trp).  
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).  
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: LA9336; AAC05711.1; -;  
DR HSSP: P00953; IDZR.  
DR InterPro: IPR002305; tRNA-synt\_1b.  
DR InterPro: IPR001412; tRNA-synt\_1.  
DR InterPro: IPR002306; tRNA-synt\_trp.  
DR Pfam: PF00579; tRNA-synt\_1b: 1.  
DR PRINTS: PR01039; TRNASYNTTRP.  
DR PROSITE: PS00178; AA\_TRNA\_LIGASE\_1; 1.  
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding.  
FT SITE 12 20 "HIGH" REGION.  
FT SITE 201 205 "KMSKS" REGION.  
FT BINDING 204 204 ATP (BY SIMILARITY).  
SQ SEQUENCE 341 AA; 38256 MW; 693C820F5A0844D1 CRC64;

Query Match 7.0%; Score 178.5; DB 1; Length 341;  
Best Local Similarity 24.2%; Pred. No. 5.2e-07;  
Matches 81; Conservative 60; Mismatches 159; Indels 35; Gaps 11;

QY 158 LYTGRGPSEAMHGHLPFTK-WLQDVENVPLVQMTDDEKYLKDLTLDAQVDAV 216  
DB 6 ILTGRPRPGK-LHIGHYVSLKNRYOLNSGDRSFIMADQALTDNARNPEKIRNSLI 64  
QY 217 ENAKIILGCGPDKNTFF--SDLYMGSSGFYKNV---KIQKHVFENQKGIFFPD 271  
DB 65 EYALDYLAAGDPLKSTILVOSQPELNLNLYLNLVLSLEKRPYKAEIKOKNEN 124  
QY 272 SDICIKISPAIOAAPSFSNSFPQIFRDTIOCLIPCAIDDPYFRMTRDVAPRIG-- 328  
DB 125 STPAGFLIYPVSGAADITN-----FKATY-----VPVGEDQLPMIEQARELVRSGNTLY 173  
QY 329 -----YKRALIHSTF--FPALQAGQTKMSADPNSSIFLTDYKQKTKYKNAKIAFSG 380  
DB 174 GKEVLVERPAVIRPKGTIGRLPCTDG-KAKMSKISGN-AIYLADADAVIKQKWSMYTDPN 231  
QY 381 RUTIEEHROFGNCQVDVSMYLFLEDDDDKLEQIRKQYTSGAMLTGELKALLEVLP 440  
DB 232 NIKVTDPGQVEEN-----YVFTYLDTFCKDTETLEBKAKYISNGIGDYKVKFLENIIDA 287  
QY 441 LIAEQARKREYDEIVKEFMPTRKLSFDFOKIAA 475  
DB 288 ELEPIRNRKKEFOKDIPREYRIKKGSEKAREVAA 322

RESULT 15  
SYCC\_YEAST  
ID SYCC\_YEAST STANDARD; PRT; 394 AA.  
AC P36421;  
DT 01-JUN-1994 (Rel. 29, Created)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last annotation update)  
DE Tyrosyl-tRNA synthetase, cytoplasmic (EC 6.1.1.1) (Tyrosyl--tRNA  
DE ligase) (TYRS).  
GN TYSI OR MGM104 OR YGR185C OR G7522.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
OX NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC MEDLINE=93286133; Pubmed=8509419;  
RA Chow C.M., Rajbhandary U.L.;  
RT "Saccharomyces cerevisiae cytoplasmic tyrosyl-tRNA synthetase gene.  
RT Isolation by complementation of a mutant *Escherichia coli* suppressor  
RT tRNA defective in aminoacylation and sequence analysis.";  
RL J. Biol. Chem. 268:12855-12863(1993).  
RN [2]  
RN SEQUENCE FROM N.A.  
RA Guan M.-X., Chen X.-J., Clark-Walker G.D.;  
RP Submitted (MAY-1993) to the EMBL/Genbank/DBJ databases.  
RC [3]  
RC STRAIN=S288C;  
RX MEDLINE=97279231; Pubmed=9133739;  
RA Arroyo J., Garcia-Gonzalez M., Garcia-Saez M.I., Sanchez-Perez M.,  
RA Nombela C.;  
RT "DNA sequence analysis of a 23,002 bp DNA fragment of the right arm  
RT of *Saccharomyces cerevisiae* chromosome VII.";  
RL Yeast 13:357-363(1997).  
RN [4]  
RP SEQUENCE OF 1-36 FROM N.A.  
RC STRAIN=B3926;  
RX MEDLINE=95087887; Pubmed=7995524;  
RA Henry N.L., Campbell A.M., Feaver W.J., Poon D., Well P.A.,  
RA Kornberg R.D.;  
RT "TritF-TAF-RNA polymerase II connection.";  
RL Genes Dev. 8:2868-2878(1994).  
CC -1- CATALYTIC ACTIVITY: ATP + L-tyrosine + tRNA(Tyr) = AMP +  
CC diphosphate + L-tyrosyl-tRNA(Tyr).  
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: L12221; AAB59329.1; -;  
DR EMBL: X71998; -: NOT ANNOTATED\_CDS.  
DR EMBL: 272970; CAA97211.1; -;  
DR EMBL: X99074; CAA67529.1; -;  
DR EMBL: U13015; AAB61641.1; -;  
DR PIR: A45999; A45999.  
DR SCD: S0003417; TYSI.  
DR InterPro: IPR002305; tRNA-synt\_1b.  
DR InterPro: IPR001412; tRNA-synt\_1.  
DR InterPro: IPR002307; tRNA-synt\_tyr.  
DR Pfam: PF00579; tRNA-synt\_1b; 1.  
DR PRINTS: PR01040; TRNASYNTTRP.  
DR PROSITE: PS00178; AA\_TRNA\_LIGASE\_1; FALSE NEG.  
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding.  
FT SITE 48 56 "HIGH" REGION.  
FT SITE 227 231 "KMSKS" REGION.  
SQ SEQUENCE 394 AA; 44020 MW; 57EBDB9BE6D054B7 CRC64;

```

CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AP000064; BAA81476.1; -
DR InterPro: IPR002305; tRNA-synt_1b.
DR InterPro: IPR001412; tRNA-synt_1.
DR InterPro: IPR002306; tRNA-synt_trp.
DR Pfam: PF00579; tRNA-synt_1b; 1.
DR PRINTS: PR01039; TRNASYNTHDRP.
DR PROSITE: PS00178; AA-TRNA_LIGASE_I; FALSE_NEG.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW Complete proteome.
FT SITE 81 89 "HIGH" REGION.
FT SITE 258 262 "KMSKS" REGION.
SQ SEQUENCE 374 AA; 42400 MW; A72635B7CA3F9189 CRC64;

Query Match 13.9%; Score 353; DB 1; Length 374;
Best Local Similarity 30.4%; Pred. No. 5.2e-21;
Matches 119; Conservative 70; Mismatches 165; Indels 38; Gaps 17;

QY 85 VDPMTVYQSSAKGIDYDKLIVRFSGSSKIDKELINIERATGQRPHHFLRGIFFSHRDMN 144
DB 8 LDPW---GAVEIKDYDRLRTFGIRPSEVL--PLRKAMEPSFLMRKGIIFGHRD 61
QY 145 QVLDAVENKKPFYLYTGRGSEAMHVGHLIPFTKWLQ-DVENVPLVIOMTDEKYLW 203
DB 62 KILEKANGENVAVLTGMPGSK-FHGHKLTVQLIYLQNGFV--FVALADEANAV 118
QY 204 KDLIDQAYGAVEN-AKDITACGFDIKT-FISDDLDMGMSGFEYNNVKIQKHVTEN 261
DB 119 RRISEEAARVAEVEYINMIALGLDPKDTGEYFQ---TNRGPRYFLIOLDFSGKYAA 174
QY 262 QVKGIFG-FTPSDICIGKISFPAIOAPFSNSFPQIFRDRDIOCLICAIIDDDPYFMT 320
DB 175 EMEAILYGLTFAKKMASLT---QADILHVLQLDYGGYR--HVVNVGADDDPHRLT 227
QY 321 RDVAPR---IGYPRPALHSTFFPALOGAOTKMSASDPNSIFLDTAKOIKTKVNKA 376
DB 228 RDLADRMAGVLEPRPASTYHKLPGLDG--KKMSSSRPDSITIFLDPREVAANKLPR-A 284
QY 377 FSGGRDITIEHRQFGNCDV-DVSEFMYITFLEDDDKLEQIRKDYTS---GAMLTGELKK 432
DB 285 LTGGRATAEQRRLGCVPEVCYVHMIDYHMLPDDGVEYKH---YTSCLRGKIILGCECKQ 341
QY 433 ALIEVLOPLIAEHOARREVTDEIVKEFMPR 464
DB 342 IAMEKLEFLEHOSRLEKAKTIAMKLVLEPR 373

RESULT 13
SYN_ARCFU STANDARD; PRT; 323 AA.
AC 029482;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Tyrosyl-tRNA synthetase (EC 6.1.1.1) (Tyrosine-tRNA ligase) (TYRS).
OS TYRS OR AF0776.
SN Archaeoglobus fulgidus.
OC Archaea; Euryarchaeota; Archaeoglobales; Archaeoglobaceae;
OC Archaeoglobus.
OX NCBI_TaxID=2234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=98049343; PubMed=9389475;

```

```

RA Kleen H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA Kechum K.A., Dodson R.J., Gwin M., Hickey E.K., Peterson J.D.,
RA Richardson D.L., Kerlavage A.R., Graham D.E., Kyprides N.C.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Lotus B.,
RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utechtack T.,
RA Cotton M.D., Spriggs T., Artlach J.P., Kaine B.P., Sykes S.M.,
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RT "The complete genome sequence of the hyperthermophilic, sulphate-
RT reducing archaeon Archaeoglobus fulgidus."
RL Nature 390:364-370(1997).
CC -1- CATALYTIC ACTIVITY: ATP + L-tyrosine + tRNA(Tyr) = AMP +
CC diphosphate + L-tyrosyl-tRNA(Tyr).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE001051; AAB90462.1; -
DR TIGR: AF0776; -
DR InterPro: IPR002305; tRNA-synt_1b.
DR InterPro: IPR001412; tRNA-synt_1.
DR InterPro: IPR002307; tRNA-synt_tyr.
DR Pfam: PF00579; tRNA-synt_1b; 1.
DR PRINTS: PR01040; TRNASYNTHTYR.
DR PROSITE: PS00178; AA-TRNA_LIGASE_I; 1.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW Complete proteome.
FT SITE 41 49 "HIGH" REGION.
FT SITE 214 218 "KMSKS" REGION.
SQ SEQUENCE 323 AA; 36616 MW; A655AEFA5116642 CRC64;

Query Match 7.5%; Score 192; DB 1; Length 323;
Best Local Similarity 23.1%; Pred. No. 3.9e-08;
Matches 78; Conservative 67; Mismatches 127; Indels 66; Gaps 15;

QY 139 SHRMNOVLDAVENKKPFYLYTGRGSEAMHVGHLIPFTKWLQVFNPLVIOMTDD 198
DB 19 TEEELROLLETKEKPR--ATVGYEPGSE-THLGHMMTVOKLMDIQE-GFEITVLADI 73
QY 199 EKYLMKDLTLDQAVGAVENAKDILAGFDINKFTFISDDLDMGMSGFEYNNVKIQKH 258
DB 74 HAYLNKGTFEEIADVADYNNKVFIALGLDSRAKFFVLSGYQ-LSRDYVLDVAKMARIT 132
QY 259 TFCNVK---GIFGFTSDICIGKISFPAIOAPFSNSFPQIFRDRDIOCL-TPCA--- 310
DB 133 TLNRRARSMDVSRKREDPMVSCMITYPLMDA-----LDIAHLGVDAVLG 177
QY 311 IDDDPYFRMTDVAAPRIGYPRPALHSTFFPALOGAOTKMSASDPNSIFLDTAKOIKT 370
DB 178 IDOKRIHMLAREN.PRLGYSSPVCPLHPPIVLGDLG--QKSSSKGN-YISRDPEEVE 234
QY 371 KVNK-HAFSG-----GRDTIEHRQFGNCDVDFSEFMYITFLEDD 410
DB 235 KIRKAYCPAGVEENPLIDIAKHILIRFGKIYVERAKRG--DVE-----YASF----- 283
QY 411 DKLEQIRKDYTSAGMLTGEKALIEVLOPLIAEHOAR 448
DB 284 ---BELAEDFRSGQLHPLDKIAVAKYIMLLEDARRK 318

RESULT 14
SYN_CLOLO STANDARD; PRT; 341 AA.
ID SYM_CLOLO

```

```

FT  NON_TER      134      134
SQ  SEQUENCE     134 AA: 15744 MW: 60E2935B7E1E344F CRC64:
Query Match      15.2%; Score 386; DB 1; Length 134;
Best Local Similarity 51.5%; Pred. No. 2.7e-24;
Matches 68; Conservative 25; Mismatches 35; Indels 4; Gaps 1;
QY  82 EDFVPMVYQVTSASAK---GIDYDKLIYFGSSKIDKELINRIERATGCRPHHFLRGIF 137
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB  3 EGRITPMDEVEVSTDEVPALIDYDKINIOFGCEKFMALADLEKISGKPAHYFFRGIV 62
QY  138 ESHRDMNOVLADYENKKPPYLYTGRGPSSSEAMHVGHLIPFTKMLQDVFNPLVYQMTD 197
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB  63 FAHRPFNLIDLIANNRPPELYTGRGPSSSKMTWHTTIFPLCKYQDAFKIRLVQITD 122
QY  198 DEKYLAKDITLD 209
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB  123 DEKFLMKSRLE 134

RESULT 11
SYN_ARCFU
ID  SYN_ARCFU      STANDARD:      PRT:      420 AA.
AC  028579;
DT  15-JUL-1998 (Rel. 36, Created)
DT  15-JUL-1998 (Rel. 36, Last sequence update)
DT  16-OCT-2001 (Rel. 40, Last annotation update)
DE  Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA ligase)
    (TRPRS5)
CN  TRPS OR AFI694.
OS  Archaeoglobus fulgidus.
OC  Archaea; Euryarchaeota; Archaeoglobales; Archaeoglobaceae;
OC  Archaeoglobus.
OX  NCBI_TaxID=2234;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX  MEDLINE=98049343; PubMed=9389475;
RA  Kleink H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA  Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.R., Peterson J.D.,
RA  Richardson D.L., Kervavage A.R., Graham D.E., Kyprides N.C.,
RA  Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA  Kitzness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA  Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA  Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Uterback T.,
RA  Cotton M.D., Spriggs T., Artiach P., Kalne B.P., Sykes S.M.,
RA  Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA  Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA  Venter J.C.;
RT  "The complete genome sequence of the hyperthermophilic, sulphate-
RT  reducing archaeon Archaeoglobus fulgidus."
RL  Nature 390:364-370(1997).
CC  -i- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +
CC  diphosphate + L-tryptophanyl-tRNA(Trp).
CC  -i- SUBCELLULAR LOCATION: Cytoplasmic.
CC  -i- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-tRNA SYNTHETASE FAMILY.
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb.ch/announce/
CC  or send an email to license@isb.ch).
CC  -----
DR  EMBL: AE000986; AAB89554.1; -
DR  TIGR: AFI694; -
DR  InterPro: IPR002305; tRNA-synt_1b.
DR  InterPro: IPR001412; tRNA-synt_1.
DR  InterPro: IPR002306; tRNA-synt_trp.
DR  Pfam: PF00579; tRNA-synt_1b; 1.
DR  PRINTS: PR01039; TRNASYNTTRP.
DR  PROSITE: PS00178; AA-tRNA_LIGASE_1; FALSE_NEG.

```

```

KM  Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KN  Complete proteome..
FT  SITE      72      80      "HIGH" REGION.
FT  SITE      308     312     "KMSK" REGION.
SQ  SEQUENCE     420 AA: 47283 MW: 9315152E2F172F24 CRC64:
Query Match      14.5%; Score 370.5; DB 1; Length 420;
Best Local Similarity 27.3%; Pred. No. 2.4e-22;
Matches 118; Conservative 60; Mismatches 166; Indels 89; Gaps 12;
QY  85 VDPVYVQVSSAGIDYDKLIYFGSSKIDKELINRIERATGCRPHHFLRGIFFSHRDMN 144
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB  3 VTPWEVEGV---IDYSKLEEFQMOPF-SEVLPEID-----NPHLMRGAIIFGRDWM 52
QY  145 OVLDAVENKKPPYLYTGRGPSSSEAMHVGHLIPFTKMLQDVFNPLVYQMTDDEKYLK 204
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB  53 RILEAMQKKEPAAVMSGFPSPG-LPHFGKMTMDLVYHQSGAKAFV-AIDMEHSHR 110
QY  205 DLTLDQAYGDAVENAKDIACGFDINKTFIFSDLDYMGMSGFYKNV-KIQKHYTFNOV 263
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB  111 GLSMEXTRELGLMLYKSIILALGLREDAYIFQS-----KSHVYKDLAFELSAEVNFSEL 164
QY  264 KQIFGFTSDCTGKISFPALQAPSPNSFPQIFRDRITQCLICALDQDYPFRMTD 323
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB  165 RAIFYGNSDTSIAKMFVTAIOAADII--HPQLSPGGRPVVPGADODPHMRLTRDL 221
QY  324 APRI-----
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB  222 AARISIFSEPEYEGVAVRSRKGAEYLSLRLEPDKIYEEHMDIFGAELERAVRKI 281
QY  328 -----GYKPDALHSTFPALQAGQTKMSASDPNSSIFLDTAKQIKTKVKNKHAFFSG 380
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB  282 EVEIGGAFATIPPSYVHRTTGLTG--KMSSSKPEYSYSLDPEEGAKKVMK-AFTGG 338
QY  381 RDTIEHRQFGNCQVDVDFMWLTFLFD-DKLDIRKDYSGAMLTBELKALILEVQ 439
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB  339 RATAEQRRLGGEPCVAFELYSPLIDSDDELNOIEACREGRLCGCKRMAAEIVK 398
QY  440 PLIAEHQARRKEY 452
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB  399 SFLKEHQEKMEAV 411

RESULT 12
SYN_AERPE
ID  SYN_AERPE      STANDARD:      PRT:      374 AA.
AC  09Y924;
DT  30-MAY-2000 (Rel. 39, Created)
DT  30-MAY-2000 (Rel. 39, Last sequence update)
DT  16-OCT-2001 (Rel. 40, Last annotation update)
DE  Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA ligase)
    (TRPRS5)
CN  TRPS OR APE2461.
OS  Aeropyrum pernix.
OC  Archaea; Crenarchaeota; Desulfurococcales; Desulfurococcaceae;
OC  Aeropyrum.
OX  NCBI_TaxID=56636;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=KI;
RX  MEDLINE=99310339; PubMed=10382966;
RA  Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
RA  Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankai A., Kosugi H.,
RA  Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
RA  Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
RA  Yamazaki J., Kushida N., Oguchi A., Aoki K.-I., Kubota K.,
RA  Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
RT  "Complete genome sequence of an aerobic hyper-thermophilic
RT  crenarchaeon, Aeropyrum pernix KI."
RL  DNA Res. 6:83-101(1999).
CC  -i- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +
CC  diphosphate + L-tryptophanyl-tRNA(Trp).
CC  -i- SUBCELLULAR LOCATION: Cytoplasmic.

```

```

Db 113 RMSEFTKELALNEYITVIALGLDPEKINIVYLSQSKYKV-----KDLALISKRTNMS 167
Qy 263 QVKGIFGFSDSCIGKISFPAIOAPSFNSFPQIFRDKT--DIOCLIFCALDODPYFMM 319
Db 168 EKKALIFGERGETINIHVAPIVQVADIL--HPQDENLSPPEKPPVAVGVGIDODPHRL 224
Qy 320 TRDVAIPR---IGYPPALHSTFFPALOGAGOTKMSASDPNSSIFLDTAKOIKTKVKNHA 376
Db 225 TRDIANKRAKEFFIPSPSYTHRFMTGLGG--KMSSSKRETAIFLITDDEKTVKKKIFS-A 281
Qy 377 FSGGDTIEHRQFG--NCDVDSFMYLTFLEDDDKLEQIRKDYTSGAMLTGELKKA 433
Db 282 KTGRETLEHKKYGGVPECVYVELFLY--HLILDKELEIYQKCRSGELTGCKCKM 339
Qy 434 LIEVQPLAEHQARKREYTDIVK 458
Db 340 AYERVEEFLDKLKERQAKELAVK 364

RESULT 9
ID SYM_METH STANDARD: PRT: 364 AA.
AC 026352:
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA ligase)
DE (TrpRS).
GN TRPS OR MTH251.
OS Methanobacterium thermoautotrophicum.
OC Archaea: Euryarchaeota; Methanobacteriales; Methanobacteriaceae;
OX Methanothermobacter.
XX NCBI_TaxID=145262;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-DELTA H:
RX MEDLINE=98037514; PubMed=9371463;
RA Smith D.R., Doucelte-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,
RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
RA Harrison D., Hoang L., Keagle P., Lumm W., Potlter B., Qiu D.,
RA Spadafora R., Vitare R., Wang Y., Wierzbowski J., Gibson R.,
RA Jivani N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,
RA McQuaglin S., Shimer G., Goyal A., Pietrovski S., Church G.M.,
RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;
RT "Complete genome sequence of Methanobacterium thermoautotrophicum
RT deltaH: functional analysis and comparative genomics.";
RL J. Bacteriol. 179:7135-7155(1997).
CC -I- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +
CC diphosphate + L-tryptophanyl-tRNA(Trp).
CC -I- SUBCELLULAR LOCATION: Cytoplasmic.
CC -I- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AE000812; AAB84757.1; -
DR InterPro: IPR002305; tRNA-synt_1b.
DR InterPro: IPR001412; tRNA-synt_1.
DR InterPro: IPR002306; tRNA-synt_1b.
DR Pfam: PF00579; tRNA-synt_1b; 1.
DR PRINTS: PRO1039; TRNASYNTTRP.
DR PROSITE: PS00178; AA_TRNA_LIGASE_I; FALSE NEG.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW Complete proteome.
FT SITE 70 "HIGH" REGION.
FT SITE 251 "KMSKS" REGION.
FT SEQUENCE 364 AA; 41301 MW; C2F348903338F61D CRC64;

```

```

Query Match 15.6%; Score 397.5; Db 1; Length 364;
Best Local Similarity 27.9%; Pred No. 1,3e-24;
Matches 107; Conservative 76; Mismatches 153; Indels 47; Gaps 14;

Qy 85 VDPWTQVOTSSAKIDYDKLIVRFGSSKIDKELINRIERATGGRPHFLRGIGFFSHRDN 144
Db 2 IDPW-----GSAR-LVEQDLENFGRPF-SEVLDEV-----DEPSWLMRGILFGHRYE 50
Qy 145 QYLDAYENKKRPYLYTGKRPSSAMHVGHLIFITFKVLQDYFNNPVLQYMTDDEKYLK 204
Db 51 RIISAMKKEDEFAVYTGMPGSR-MHIGKMTVDQLRW-YDMGAEIIFPIDMEAYSAK 108
Qy 205 DLTLOAVGDAVEN-AKDIACGFDINK-----TFIRSDIDYMGSSGFKNYVK 253
Db 109 GYDFEDSRRIALIEEYIAGTIALGLDLEKDNHIVYLOSEMLMEDLAY----- 156
Qy 254 ICKHVTFNQKCIFFPTSDCIGKISFPAIOAPSFNSFPQIFRDRDIOCLPCAIQD 313
Db 157 LAGKVNENELRALYGFSTGSTMAMHYAPILQVSDILHPQDLBGLGPR--PYIVPGPDQ 213
Qy 314 DPEFRMTDVAIRI-----GYPPALHSTFFPALOGAGOTKMSASDPNSSIFLDTAKOIK 369
Db 214 DPHIRLTRDIARFRDRYGFIIPSSSYTHRFMTGLTG--KMSNRPKSAIFLSDPPEAE 271
Qy 370 TVKNHAFSGGDTIEHRQFGNCDVDSFMYLTFLEL-DDDKLEQIRKDYTSGAMLTG 428
Db 272 AKI-RNAKGTRETLKEQRELGGVPECCITETLLKHMGSQSRLEIYESCRNGLMG 330
Qy 429 ELKALIEVQPLAEHQARKRE 451
Db 331 ECKNNTAFIRKFFELSVKREK 353

RESULT 10
ID SYM_ENCCU STANDARD: PRT: 134 AA.
AC 096771:
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA ligase)
DE (TrpRS) (Fragment).
OS Encephalitozoon cuniculi.
OC Eukaryota; Microsporidia; Unikaryonidae; Encephalitozoon.
OX NCBI_TaxID=6035;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98277683; PubMed=9615449;
RA Peyretallade E., Broussolle V., Peyret P., Metenler G., Gouy M.,
RA Vivares C.P.;
RT "Microsporidia, amitochondrial protists, possess a 70-kDa heat shock
RT protein gene of mitochondrial evolutionary origin.";
RL Mol. Biol. Evol. 15:683-689(1998).
CC -I- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +
CC diphosphate + L-tryptophanyl-tRNA(Trp).
CC -I- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -I- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AJ012470; CAAL0034.1; -
DR HSP: P00952; ITC.
DR InterPro: IPR001412; tRNA-synt_1.
DR PROSITE: PS00178; AA_TRNA_LIGASE_I; 1.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding.
FT SITE 89 "HIGH" REGION.

```

```
STW_PYRAB
ID SYM_PYRAB STANDARD: PRT: 385 AA.
AC Q9UY11;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE tryptophanyl-tRNA synthetase (EC 6.1.1.2) (tryptophan--tRNA ligase)
DE (TrpRS).
GN TRPS OR PAB111.
OS Pyrococcus abyssi.
OC Archaea: Euryarchaeota: Thermococcales; Thermococcaceae; Pyrococcus.
OX NCBI_TaxID=29292;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ORSAY;
RT "Pyrococcus abyssi genome sequence: insights into archaeal chromosome
structure and evolution.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +
diphosphate + L-tryptophanyl-tRNA(Trp).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@sib-sib.ch).
CC -----
DR EMBL; AJ248288; CAB50601.1; -
DR InterPro; IPR002305; tRNA-synt_1b.
DR InterPro; IPR001412; tRNA-synt_1.
DR InterPro; IPR002306; tRNA-synt_1trp.
DR Pfam; PF00579; tRNA-synt_1b; 1.
DR PRINTS; PR01039; TRNASYNTHTRP.
DR Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
Complete proteome.
KW Complete proteome.
FT SITE 82 "HIGH" REGION.
FT SITE 253 "KMSKS" REGION.
FT SITE 257 "KMSKS" REGION.
SQ SEQUENCE 385 AA; 45100 MW; 4C29D01414976B12 CRC64;

Query Match 31.5%; Score 803; DB 1; Length 385;
Best Local Similarity 45.3%; Pred. No. 3.1e-57;
Matches 173; Conservative 63; Mismatches 120; Indels 26; Gaps 9;

OY 82 EDF-VDPMTVQTSSAKGIDYDKLIVRGSSKIDKELINIERATQORPHHFLRGIFFSH 140
DB 3 EDFKVTPEVEGEV---VDYNKLIEHFGSPLETEELLEKTATLTKSELPLFFRRKFFFSH 58
OY 141 RDMNOVLAAEKKRPFYLYTGRGSSSEAMVGHILPIFTKLDVFNVPVLYIOMTDEX 200
DB 59 RQYDVLDVDEGRGFFLYTGRGPG-PMHIGHIIPFATKWLQKFGVNLTIQITLDEK 117
OY 201 YLMKD-LTLDAQVGDVAENAKDIIAGCFDINKFTFSDIDYMGSSGCVKNVAKIOKHT 259
DB 118 FLFKENLIFEDRKHAYENIIDIIVGDPDKTIFQNSEF---TKIYEMAIPLAKKIN 173
OY 260 FNOVVGIGFTDSDICIGISPPAIDAPSEFSPQIFRDRDTQICLIPCAIDDPYFRM 319
DB 174 FSMARAVGFTFOSKIGMIFPAIDAPTF-----FEKR---RCLIPALAIODDPYML 223
OY 320 TRDVPVIRGYPRLHSHFPALGAGTQKMSASDPNSSIFLTDAKQIKTPVKNHARG 379
DB 224 QDRPESIGYKTTAIIHSKVPSTLSGKMSASRPETALITLDSPEDEKVMKFAALTG 283
OY 380 GRDTIEEHROFGNCDVDSFMYLTFLEDDDKBOIRKYD---TSGAMLTGELKKALIE 436
DB 284 GRPTLKQREKGEKPEKCVAFKMLIFFEEDDK--KLKERYACAKNGELTGCECKRYLIS 341
OY 437 VLQPLIAEHQARRKEVTDIYK 458
```

```
DB 342 KIQEFLEKHEKRRKRAKQIEK 363
:| : ||| ||| :| |
RESULT 8
STW_METJA
ID SYM_METJA STANDARD: PRT: 370 AA.
AC Q58810;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE tryptophanyl-tRNA synthetase (EC 6.1.1.2) (tryptophan--tRNA ligase)
DE (TrpRS).
GN TRPS OR MJ1415.
OS Methanococcus jannaschii.
OC Archaea: Euryarchaeota: Methanococcales; Methanococcaceae;
OC Methanococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reisch C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geophagen N.S.M., Weidman J.F., Fuhmann J.L., Nguyen D.,
RA Uitterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Kleck H.-P., Fraser K.M., Smith H.O., Woese C.R., Venter J.C.;
RA "Complete genome sequence of the methanogenic archaeon, Methanococcus
jannaschii.";
RT Science 273:1058-1073(1996).
RL Science 273:1058-1073(1996).
CC -1- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +
diphosphate + L-tryptophanyl-tRNA(Trp).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@sib-sib.ch).
CC -----
DR EMBL; U67582; AAB99425.1; -
DR TIGR; MJ1415; -
DR InterPro; IPR002305; tRNA-synt_1b.
DR InterPro; IPR001412; tRNA-synt_1.
DR InterPro; IPR002306; tRNA-synt_1trp.
DR Pfam; PF00579; tRNA-synt_1b; 1.
DR PRINTS; PR01039; TRNASYNTHTRP.
DR PROSITE; PS00178; AA-TRNA_LIGASE_I; FALSE NEG.
DR Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
Complete proteome.
KW Complete proteome.
FT SITE 75 "HIGH" REGION.
FT SITE 255 "KMSKS" REGION.
FT SITE 259 "KMSKS" REGION.
SQ SEQUENCE 370 AA; 42660 MW; E6C71107CF82B59D CRC64;

Query Match 16.1%; Score 409.5; DB 1; Length 370;
Best Local Similarity 30.9%; Pred. No. 1.4e-25;
Matches 119; Conservative 72; Mismatches 153; Indels 41; Gaps 16;

OY 87 PWTVQTSSAKGIDYDKLIVRGSSKIDKELINIERATQORPHHFLRGIFFSHDMNOV 146
DB 8 PW-ETPAV--IDYKKTQEGVGRPIVDLQDKEE-----HHFRNNIILGHADFERI 57
OY 147 LDVAENKRPFYLYTGRGSSSEAMVGHILPIFTKLDVFNVPVLYIOMTDEKYL 203
DB 58 VDAIKNNKEFNAVSGMMSGK-MHGKHKMAYVDLKFQYKYNININIP-----ADLEAFAWA 112
OY 204 KDULTIDQAYGDV-ENAKDIIAGCFDINKFTFSDIDYMGSSGCVKNVAKI-QKHVTFN 261
```

```

      RT Including the Tyl-H3 retrotransposon, the sufl(+) frameshift
      RT suppressor gene for tRNA-Gly, the yeast transfer RNA-Thr-1a and a
      RT delta element."
      RL Yeast 11:1069-1075(1995).
      RN [2]
      RP FUNCTION.
      RX MEDLINE=97197969; Pubmed=9046085;
      RA John T.R., Ghosh M., Johnson J.D.;
      RT "Identification and expression of the Saccharomyces cerevisiae
      RT cytoplasmic tryptophanyl-tRNA synthetase gene.";
      RL Yeast 13:37-41(1997).
      CC -1- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +
      CC diphosphate + L-tryptophanyl-tRNA(Trp).
      CC -1- SUBUNIT: HOMODIMER.
      CC -1- SUPCELLULAR LOCATION: Cytoplasmic.
      CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-tRNA SYNTHETASE FAMILY.
      CC -----
      CC This SWISS-PROT entry is copyright. It is produced through a collaboration
      CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
      CC the European Bioinformatics Institute. There are no restrictions on its
      CC use by non-profit institutions as long as its content is in no way
      CC modified and this statement is not removed. Usage by and for commercial
      CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
      CC or send an email to license@isb-sib.ch).
      CC -----
      DR EMBL: Z48149; CAAB8164.1; -.
      DR EMBL: Z74839; CAA99110.1; -.
      DR SGD: S0005457; WRS1.
      DR InterPro: IPR002305; tRNA-synt_1b.
      DR InterPro: IPR001412; tRNA-synt_1.
      DR InterPro: IPR002306; tRNA-synt_1b.
      DR Pfam: PF00579; tRNA-synt_1b; 1.
      DR PRINTS: PR01039; TRNASYNTHTRP.
      DR PROSITE: PS00178; AA.tRNA.LIGASE.1; 1.
      KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding.
      FT SITE 111 120 "HIGH" REGION.
      FT SITE 295 299 "KMSK" REGION.
      SQ SEQUENCE 432 AA; 49350 MW; C408E169737E9736 CRC64;

Query Match 45.7%; Score 1163; DB 1; Length 432;
Best Local Similarity 54.8%; Pred. No. 3,7e-86;
Matches 217; Conservative 71; Mismatches 100; Indels 8; Gaps 3;

QY 77 ATAEEDPEVDPMV-----QTSAGKIDYDKLIVFGSSKIDKELINRIERATGQRPHE 131
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 19 STDVKEOVYIPWDEGGVDEQGRAONIDYDKLIFQFTKPYNEETLKRKQVATGREPHNF 78

QY 132 LRRIEFSHDMONVLDAYENKKRPFYLYTGTGSSSAMHVGHLIRPFTKMLQOVFNPL 191
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 79 LRKLEPFSERDFKILLDYEGKRFELYLTGTGSPSSNHLHMLPFTKMLEVDVPL 138

QY 192 VIQMTDEKYLK-DLTLDQAYGDAVENAKDIILGCGEDINKTFIFSDLDYMGSSGEYKN 250
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 139 VIELTDEKELFKELKLTINDYKNFARENAKIILAVGDPKTFIFSDLYWG--GAFEY 196

QY 251 VVKIQKHVTENQVGIQGFITDSDCIKTSIPAIDAAFSFSFPQIIRDRIDIOCLIPCA 310
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 197 VVRSRRIITSGTAAVAGFNDSDCIGFHFHSIDIAFAFSSFPFNVLGLPDKTCLLPICA 256

QY 311 IDDPYPRMRIDVAPRIGYKPKALLHSITFPFALOGAOTKMSASPNISIEFLDTAKOIKT 370
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 257 IDDPYPRVRCRDVADKIKYSRPAHLHSRFPFALOGSTTKMSASDPTALAFMTDTPKQOK 316

QY 371 KVNKAHSSGGRDITIEHRQPGNCVDVSEMYITLFFLEDDDKLQIRKDYTSGMLTGL 430
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 317 KINKAASGGGVASDLTIRELGGNDVVAAYOYLTFEFDKDVFLKECYDKYKSGELLGEM 376

QY 431 KKALIEVLQPLIAEHQARREVTDEIKYKEFMPKRL 466
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 377 KKLCIEFLQEVKAFQERRAOVDEETLDKEVWPKRL 412

```



```

QY 180 TKMLQDVNPLVIGMTDEKYLKMDLTLDQAYGDAVENAKDIACGFIDNKFIFSDLD 239
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 184 TKMLQDVNPLVIGMTDEKYLKMDLTLDQAYGDAVENAKDIACGFIDNKFIFSDLE 243
QY 240 YMGSSGFYKRVNVIQKHVTEVNOVKGIFGFTDSDCIGKISFPAIDAPSFNSPQIFRD 299
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 244 YMGSGPGFYKRVNVIQKHVTEVNOVKGIFGFTDSDCIGKISFPAIDAPSFNSPQIFRD 303
QY 300 RTDIOCLIPCAIDDPYRPMTRDVAIRIGYPRKALHSTFFPALOGAQTMSASPNSSI 359
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 304 RTDIOCLIPCAIDDPYRPMTRDVAIRIGYPRKALHSTFFPALOGAQTMSASPNSSI 363
QY 360 FLTDIAKQIKTRVKNHAFSGGRDTEEHROFGNCDVDVSEFMYLTFELEDDEKLEQIRRD 419
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 364 FLTDIAKQIKTRVKNHAFSGGRDTEEHROFGNCDVDVSEFMYLTFELEDDEKLEQIRRD 423
QY 420 YTGSMALTGELKALIEVLQPLIAEHQARKREVDTEIYKEFMPKRLSPDFQ 471
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 424 YTGSMALTGELKALIEVLQPLIAEHQARKREVDTEIYKEFMPKRLSPDFHQ 475

RESULT 4
SYN_RABIT STANDARD: PRT: 475 AA.
AC P23612; Q28607;
DT 01-NOV-1991 (Rel. 20, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE tryptophanyl-tRNA synthetase (EC 6.1.1.2) (tryptophan--tRNA ligase)
GN (TPRS).
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90239043; PubMed=2185472;
RA Lee C.C., Craigen W.J., Muzny D.M., Harlow E., Caskey C.T.;
RT "Cloning and expression of a mammalian peptide chain release factor
RT with sequence similarity to tryptophanyl-tRNA synthetases.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:3508-3512(1990).
RN [2]
RP REVISIONS TO 169-174 AND 227-228, AND FUNCTION.
RX MEDLINE=94009008; PubMed=8404867;
RA Frolova L.Y., Dalphin M.E., Justesen J., Powell R.J., Druegeon G.,
RA McCaughan K.R., Kisselev L.L., Tate W.P., Haenni A.-L.;
RT "Mammalian polypeptide chain release factor and tryptophanyl-tRNA
RT synthetase are distinct proteins.";
RL EMBO J. 12:4013-4019(1993).
CC -1 CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +
CC diphosphate + L-tryptophanyl-tRNA(Trp).
CC -1 SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1 SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
CC -1 SIMILARITY: CONTAINS 1 "WHEP-TRS" DOMAIN.
CC -1 CAUTION: WAS ORIGINALLY (REF.1) THOUGHT TO BE AN EUKARYOTIC
CC RELEASE FACTOR (ERF).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: M33460; AAA31246.1; ALT_SEQ.
CC EMBL: U02595; AAB60257.1; -.
CC PIR: A35904; YWRBPR.
CC InterPro: IPR0000738; WHEP-TRS.
CC InterPro: IPR000305; tRNA-syn_1b.
CC InterPro: IPR001412; tRNA-syn_1.

```

```

DR InterPro: IPR002306; tRNA-syn_1-trp.
DR Pfam: PF00579; tRNA-syn_1b; 1.
DR Pfam: PF00458; WHEP-TRS; 1.
DR PRINTS: PR01039; TRNASYNTHTRP.
DR PROSITE: PS00176; AA_TRNA_LIGASE_1; 1.
DR PROSITE: PS00762; WHEP-TRS; 1.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding.
FT DOMAIN 23 68
FT SITE 168 177 WHEP-TRS.
FT SITE 353 357 "HIGH" REGION.
SO SEQUENCE 475 AA; 53799 MW; 33BC9E718FF45DC4 CRC64;

Query Match 86.1%; Score 2192; DB 1; Length 475;
Best Local Similarity 89.0%; Pred. No. 5,5e-169;
Matches 414; Conservative 24; Mismatches 27; Indels 0; Gaps 0;

QY 7 ASLELFNSIATOGELVSLKAGNASKDEIDSAVMYLSKMSYAAAGEDYKADCPGN 66
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 11 ASPOELFSSIAOGELVSLKARKAPKEIDSAVMYLSKMSYAAAGEDYKADCPGN 70
QY 67 PAPTSNHGPDATAEADDPVDPWTOTSSAKGIDYDKLIVRFSSKIDELIRIRATGQ 126
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 71 STPDSHGDPDAVDKEDVDVDPWTOTSSAKGIDYDKLIVRFSSKIDELIRIRATGQ 130
QY 127 RPHNPLRKGIFPSHRDMNOVDAYENKKPFYLYTGRGSSSEAHVGHILPFIETKMLQDV 186
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 131 RPHRFLRKGIFPSHRDMNOVDAYENKKPFYLYTGRGSSSEAHVGHILPFIETKMLQDV 190
QY 187 FNVPLVIGMTDEKYLKMDLTLDQAYGDAVENAKDIACGFIDNKFIFSDLDYMGSSG 246
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 191 FNVPLVIGMTDEKYLKMDLTLDQAYGDAVENAKDIACGFIDNKFIFSDLDYMGSSG 250
QY 247 FKNVNVKIOKHVTEVNOVKGIFGFTDSDCIGKISFPAIDAPSFNSPQIFRDRTDIOCL 306
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 251 FKNVNVKIOKHVTEVNOVKGIFGFTDSDCIGKISFPAIDAPSFNSPQIFRDRTDIOCL 310
QY 307 IPCAIDDPYRPMTRDVAIRIGYPRKALHSTFFPALOGAQTMSASPNSSIPLTDIAK 366
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 311 IPCAIDDPYRPMTRDVAIRIGYPRKALHSTFFPALOGAQTMSASPNSSIPLTDIAK 370
QY 367 QIKTVNKNHAFSGGRDTEEHROFGNCDVDVSEFMYLTFELEDDEKLEQIRDYSGAML 426
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 371 QIKTVNKNHAFSGGRDTEEHROFGNCDVDVSEFMYLTFELEDDEKLEQIRDYSGAML 430
QY 427 TGEELKALIEVLQPLIAEHQARKREVDTEIYKEFMPKRLSPDFQ 471
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 431 TGEELKALIEVLQPLIAEHQARKREVDTEIYKEFMPKRLSPDFQ 475

RESULT 5
SYN_SCHPO STANDARD: PRT: 395 AA.
AC Q09692;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Putative tryptophanyl-tRNA synthetase (EC 6.1.1.2) (tryptophan--tRNA
DE ligase) (TPRS).
GN SPAC2F.13C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomyces.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Gentles S., Churcher C.M., Barrell B.G., Rajandream M.A., Walsh S.V.;
RL Submitted (JUL-1995) to the EMBL/Genbank/DBJ databases.
CC -1 CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +
CC diphosphate + L-tryptophanyl-tRNA(Trp).
CC -1 SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
CC -----

```

```
CC -1- SIMILARITY: BELONGS TO CLASS-1 AMINOACYL-TRNA SYNTHETASE FAMILY.
CC -1- SIMILARITY: CONTAINS 1 "WHEP-TRS" DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X53918; CAA37872.1; -
CC EMBL: X52113; CAA36356.1; -
CC EMBL: M74074; AAA30799.1; -
CC PIR: A40279; YWBO.
CC PIR: S14540; S14540.
CC InterPro: IPR000738; WHEP-TRS.
CC InterPro: IPR002305; tRNA-synt_1b.
CC InterPro: IPR001412; tRNA-synt_1.
CC InterPro: IPR002306; tRNA-synt_trp.
CC Pfam: PF00579; tRNA-synt_1b; 1.
CC Pfam: PF00458; WHEP-TRS; 1.
CC PRINTS: PR01039; TRNASYNTHTRP.
CC PROSITE: PS00178; AA_TRNA_LIGASE_I; 1.
CC PROSITE: PS00762; WHEP-TRS; 1.
CC KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding.
CC FT DOMAIN 24 69 WHEP-TRS.
CC FT SITE 1 117 DISPENSABLE TO THE CATALYTIC ACTIVITY.
CC FT SITE 169 178 "HIGH" REGION.
CC FT SITE 353 357 "KMSKS" REGION.
CC FT CONFLICT 17 17 L -> M (IN REF. 2).
CC FT SEQUENCE 475 AA; 53729 MW; F7E531750137EB32 CRC64;

Query Match 89.8%; Score 2286.5; DB 1; Length 475;
Best Local Similarity 92.6%; Pred. No. 1.3e-176;
Matches 438; Conservative 12; Mismatches 20; Indels 3; Gaps 2;

QY 1 MNSERP--ASLLEFNSTATOGELVRSKAGNASKDEIDSAVKMLVSLKMSYKAAAGEY 58
DB 4 MSNGSGGGSPLELFHSTIAOGELVRDLKARNAKDEIDSAVKMLSLKTSKATGEY 63
QY 59 KADCPGNPAPTSNHPDPATEAEEDFVDPWTVQTSAGKIDYDKLIVRGSSKIDKELIN 118
DB 64 KYDCPPGDPAPESGEGDLATEDEDFVDPWTVQTSAGKIDYDKLIVRGSSKIDKELIN 123
QY 119 IERATGGRPHRLRGRGIFFSHRDMNQVLDAVENKKPFYLYTGRGPSSSEAMHGLIPFI 178
DB 124 IERATGGRPHRLRGRGIFFSHRDMNQVLDAVENKKPFYLYTGRGPSSSEAMHGLIPFI 183
QY 179 FTKMLQDVFNPLVQMTDDEKYLKMDLTDQAYGDAVENAKDIACGFDIKTFIFSDL 238
DB 184 FTKMLQDVFNPLVQMTDDEKYLKMDLTDQAYGDAVENAKD-ITCGFDINKTFIFSDL 242
QY 239 DYMGSSGFFYKNVYKIQKHVTFNQYKIGFTDSDICIKISPPAIOAAPFSNSPQJIFR 298
DB 243 DYMGSSGFFYKNVYKIQKHVTFNQYKIGFTDSDICIKISPPAIOAAPFSNSPQJIFR 302
QY 299 DRTDQCLIPCAIODDPFRFMTRDVAPRIQYKPKPALHSTFPFPAIOGQOTKKSADPNSS 358
DB 303 DRTDQCLIPCAIODDPFRFMTRDVAPRIQYKPKPALHSTFPFPAIOGQOTKKSADPNSS 362
QY 359 IFLDPAKQIKTKYKNAHAFSGRDTIEEHROFGNCDVDVSPMYLTFLEDDDKLEQJIR 418
DB 363 IFLDPAKQIKTKYKNAHAFSGRDTIEEHROFGNCDVDVSPMYLTFLEDDDKLEQJIR 422
QY 419 DYTSGAMLTGELKALIEVLOPLIAHQARREYTDIVKEFMPRKLSDFQ 471
DB 423 DYTSGAMLTGELKALIEVLOPLIAHQARREYTDIVKEFMPRKLSDFQ 475
```

```
AC P32921;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE tryptophanyl-tRNA synthetase (EC 6.1.1.2) (tryptophan--tRNA ligase)
DE (trpRS).
GN WARS OR WRS.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95018226; PubMed=7932716;
RA Pajot B., Sarger C., Bonnet J., Garret M.;
RT "An alternative splicing modifies the C-terminal end of tryptophanyl-
RT tRNA synthetase in murine embryonic stem cells.";
RL J. Mol. Biol. 242:599-603(1994).
CC -1- CATALYTIC ACTIVITY: Arp + L-tryptophan + tRNA(Trp) = AMP +
CC dihydrophosphate + L-tryptophanyl-tRNA(Trp).
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1- ALTERNATIVE PRODUCTS: TWO FORMS OF THIS PROTEIN ARE PRODUCED BY
CC ALTERNATIVE SPLICING. A ISOFORM OF 475 RESIDUES, FOUND IN MOST
CC TISSUES AND A C-TERMINALLY EXTENDED FORM OF 481 RESIDUES FOUND IN
CC EMBRYONIC STEM CELLS.
CC -1- SIMILARITY: BELONGS TO CLASS-1 AMINOACYL-TRNA SYNTHETASE FAMILY.
CC -1- SIMILARITY: CONTAINS 1 "WHEP-TRS" DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X69656; CAA49347.1; -
CC EMBL: X69657; CAA49348.1; -
CC PIR: S31461; S31461.
CC PIR: S31462; S31462.
CC MGD: MGI:104630; Wars.
CC InterPro: IPR000738; WHEP-TRS.
CC InterPro: IPR002305; tRNA-synt_1b.
CC InterPro: IPR001412; tRNA-synt_1.
CC InterPro: IPR002306; tRNA-synt_trp.
CC Pfam: PF00579; tRNA-synt_1b; 1.
CC Pfam: PF00458; WHEP-TRS; 1.
CC PRINTS: PR01039; TRNASYNTHTRP.
CC PROSITE: PS00178; AA_TRNA_LIGASE_I; 1.
CC PROSITE: PS00762; WHEP-TRS; 1.
CC KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
CC Alternative splicing.
CC FT DOMAIN 23 68 WHEP-TRS.
CC FT SITE 168 177 "HIGH" REGION.
CC FT SITE 353 357 "KMSKS" REGION.
CC FT VASPLIC 476 481 MISSING (IN MAJOR ISOFORM).
CC FT SEQUENCE 481 AA; 54282 MW; B05A452C08074F52 CRC64;

Query Match 86.7%; Score 2208.5; DB 1; Length 481;
Best Local Similarity 89.0%; Pred. No. 2.6e-170;
Matches 420; Conservative 24; Mismatches 27; Indels 1; Gaps 1;

QY 1 MNSERP--ASLLEFNSTATOGELVRSKAGNASKDEIDSAVKMLVSLKMSYKAAAGEY 59
DB 4 MSNGSGGGSPLELFHSTIAOGELVRSKAGNAPKDEIDSAVKMLSLKMSYKAAAGEYK 63
QY 60 ADCPPGNPAPTSNHPDPATEAEEDFVDPWTVQTSAGKIDYDKLIVRGSSKIDKELIN 119
DB 64 ADCPPGNPAPTSNHPDPATEAEEDFVDPWTVQTSAGKIDYDKLIVRGSSKIDKELIN 123
QY 120 IERATGGRPHRLRGRGIFFSHRDMNQVLDAVENKKPFYLYTGRGPSSSEAMHGLIPFI 179
DB 124 IERATGGRPHRLRGRGIFFSHRDMNQVLDAVENKKPFYLYTGRGPSSSEAMHGLIPFI 183
```

RA Frolova L.Y., Grigorieva A.Y., Sudomolina M.A., Kisseliev L.L.;  
 RT "The human gene encoding tryptophanyl-tRNA synthetase: interferon-  
 RT response elements and exon-intron organization.";  
 RL Gene 128:237-245(1993).  
 RN [7]  
 RP SEQUENCE OF 265-276; 278-296; 299-317 AND 350-365.  
 RC TISSUE=keratinocytes;  
 RX MEDLINE=93162043; PubMed=1286667;  
 RA Rasmussen H.H., van Damme J., Puype M., Gesser B., Celis J.E.,  
 RA Vandekerckhove J.;  
 RT "Microsequences of 145 proteins recorded in the two-dimensional gel  
 RT protein database of normal human epidermal keratinocytes.";  
 RL Electrophoresis 13:960-969(1992).  
 RN [8]  
 RP FUNCTION.  
 RX MEDLINE=92225128; PubMed=1373391;  
 RA Bange F.-C., Flohr T., Buwilt U., Boeltger E.C.;  
 RT "An interferon-induced protein with release factor activity is a  
 RT tryptophanyl-tRNA synthetase.";  
 RL FEBS Lett. 300:162-166(1992).  
 CC -I- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +  
 CC diphosphate + L-tryptophanyl-tRNA(Trp).  
 CC -I- SUBUNIT: HOMODIMER.  
 CC -I- INDUCTION: BY INTERFERON GAMMA.  
 CC -I- SIMILARITY: BELONGS TO CLASS-I "WHEP-TRS" DOMAIN.  
 CC -I- SIMILARITY: CONTAINS 1 "WHEP-TRS" DOMAIN.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: M7804; AAA67324.1; -  
 DR EMBL: X59892; CAA42545.1; -  
 DR EMBL: M61715; AAA61298.1; -  
 DR EMBL: X62570; CAA44450.1; -  
 DR EMBL: S82950; AAB39381.1; -  
 DR EMBL: X67920; CAB94198.1; -  
 DR EMBL: X67921; CAB94198.1; JOINED.  
 DR EMBL: X67922; CAB94198.1; JOINED.  
 DR EMBL: X67923; CAB94199.1; -  
 DR EMBL: X67924; CAB94199.1; JOINED.  
 DR EMBL: X67925; CAB94199.1; JOINED.  
 DR EMBL: X67926; CAB94199.1; JOINED.  
 DR EMBL: X67927; CAB94199.1; JOINED.  
 DR EMBL: X67928; CAB94199.1; JOINED.  
 DR PIR: A41706; A41706.  
 DR PIR: A41633; A41633.  
 DR PIR: JH0533; JH0533.  
 DR PIR: S19246; S19246.  
 DR Aarhus/Ghent-2DPAGE; 3524; IEF.  
 DR PHC1-2DPAGE; P23381; -  
 DR MIM: 191050; -  
 DR InterPro: IPR000738; WHEP-TRS.  
 DR InterPro: IPR002305; tRNA-synt\_1b.  
 DR InterPro: IPR001412; tRNA-synt\_1.  
 DR InterPro: IPR002306; tRNA-synt\_trp.  
 DR Pfam: PF00579; tRNA-synt\_1b.1.  
 DR Pfam: PF00458; WHEP-TRS.1.  
 DR PRINTS: PRO1039; TRNASYNTTRP.  
 DR PROSITE: PS00178; AA-TRNA\_LIGASE\_I; 1.  
 DR PROSITE: PS00762; WHEP-TRS; 1.  
 DR Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding.  
 KM DOMAIN  
 FT 19 64 WHEP-TRS.  
 FT SITE 164 173 "HIGH" REGION.  
 FT SITE 349 353 "KMSKS" REGION.  
 FT CONFLICT 213 214 SY -> GD (IN REF. 3).  
 FT CONFLICT 424 424 A -> R (IN REF. 4).  
 FT SEQUENCE 471 AA: 53165 MW: E96344449053A0D0 CRC64;

Query Match 96.3%; Score 2454; DB 1; Length 471;  
 Best Local Similarity 99.6%; Pred. No. 4,3e-190;  
 Matches 469; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 MNSEPASLLELPNSTATGELVRSIKAGNSKDEIDSAVKMLVSLKMSYKAAGEDYKA 60  
 DB 1 MNSEPASLLELPNSTATGELVRSIKAGNSKDEIDSAVKMLVSLKMSYKAAGEDYKA 60  
 QY 61 DCPGPAFTSNHGPATAEEDFVDPWTVOTSSAKGIDYDKLTIRFGSKIDKELINT 120  
 DB 61 DCPGPAFTSNHGPATAEEDFVDPWTVOTSSAKGIDYDKLTIRFGSKIDKELINT 120  
 QY 121 ERATGGRPHPLRGIFSFSDRMNOYLDAVENKKPFYLYTGGRPSSEAHVGHLPFT 180  
 DB 121 ERATGGRPHPLRGIFSFSDRMNOYLDAVENKKPFYLYTGGRPSSEAHVGHLPFT 180  
 QY 181 KMLQVFNPLVYQMTDEKTYLMDLQAYGDAVENAKDILACGFDINKFTFSDLDY 240  
 DB 181 KMLQVFNPLVYQMTDEKTYLMDLQAYGDAVENAKDILACGFDINKFTFSDLDY 240  
 QY 241 MGNSSGFYKNVYKIOKHTFNOVKGIFFGFTSDCIKISFPALIOAAPSNSFPQIFRDR 300  
 DB 241 MGNSSGFYKNVYKIOKHTFNOVKGIFFGFTSDCIKISFPALIOAAPSNSFPQIFRDR 300  
 QY 301 TDICLIPCAIDODPYFRMTDVAPRIGYKRPALLHSTFPALOGAQTKMSADPNSSIF 360  
 DB 301 TDICLIPCAIDODPYFRMTDVAPRIGYKRPALLHSTFPALOGAQTKMSADPNSSIF 360  
 QY 361 LVDTAQKITKYNKHAFFSGRPTIEHKGFGNCVDVDFMLTFFLEDKDLQIRKDY 420  
 DB 361 LVDTAQKITKYNKHAFFSGRPTIEHKGFGNCVDVDFMLTFFLEDKDLQIRKDY 420  
 QY 421 TSGAMLTGELKALLLEVLOPLAEHQARKKVTDEIVKFFMTPRKLSFDFQ 471  
 DB 421 TSGAMLTGELKALLLEVLOPLAEHQARKKVTDEIVKFFMTPRKLSFDFQ 471  
 RESULT 2  
 SYW\_BOVIN STANDARD; PRT; 475 AA.  
 ID SYW\_BOVIN  
 AC P17248;  
 DT 01-AUG-1990 (Rel. 15, Created)  
 DT 01-MAY-1992 (Rel. 22, Last sequence update)  
 DT 01-FEB-1996 (Rel. 33, Last annotation update)  
 DE Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA ligase)  
 DE (TrpRS).  
 GN WARS.  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Bovinae; Bos.  
 ON NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RC TISSUE=Retina.  
 RX MEDLINE=91329348; PubMed=1907847;  
 RA Garret M., Pajot B., Trezeguet V., Labouesse J., Merle M.,  
 RA Garret J.-C., Benedetto J.-P., Sallatranque M.-L., Alterio J.,  
 RA Gueguen M., Sarger C., Labouesse B., Bonnet J.;  
 RT "A mammalian tryptophanyl-tRNA synthetase shows little homology to  
 RT prokaryotic synthetases but near identity with mammalian peptide  
 RT chain release factor.";  
 RL Biochemistry 30:7809-7817(1991).  
 RN [2]  
 RP SEQUENCE OF 17-475 FROM N.A.  
 RC TISSUE=Pancreas;  
 RA Garret M., Trezeguet V., Pajot B., Gandar J.-C., Merle M.,  
 RA Gueguen M., Benedetto J.-P., Sarger C., Alterio J., la Bounessec B.,  
 RA Labouesse J., Bonnet J.;  
 RL Submitted (MAR-1990) to the EMBL/GenBank/DBJ databases.  
 CC -I- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +  
 CC diphosphate + L-tryptophanyl-tRNA(Trp).  
 CC -I- SUBUNIT: HOMODIMER.

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 24, 2002, 12:49:40 ; Search time 8.87395 Seconds  
(without alignments)  
2111.829 Million cell updates/sec

Title: US-09-813-718-10  
Perfect score: 2547  
Sequence: 1 MPNSEPASLLELFENSIAIQQ.....KLSFDQKLAALHHNNHH 484

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues  
Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_40:\*  
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	2454	96.3	471 1	SYW_HUMAN
2	2286.5	89.8	475 1	SYW_BOVIN
3	2208.5	86.7	481 1	SYW_MOUSE
4	2192	86.1	475 1	SYW_RABIT
5	1210	47.5	395 1	SYW_SCHPO
6	1163	45.7	432 1	SYWC_YEAST
7	803	31.5	385 1	SYW_PIRAB
8	409.5	16.1	370 1	SYW_METUA
9	397.5	15.6	364 1	SYW_METTH
10	386	15.2	134 1	SYW_ENCCU
11	370.5	14.5	420 1	SYW_ARCFU
12	353	13.9	374 1	SYW_AERPE
13	192	7.5	323 1	SYW_ARCFU
14	178.5	7.0	341 1	SYW_CLOLO
15	172	6.8	394 1	SYWC_YEAST
16	170.5	6.7	366 1	SYW_SULSO
17	169	6.6	395 1	SYW_AOUAE
18	166.5	6.5	346 1	SYW_CHLTR
19	163.5	6.4	353 1	SYW_BORBU
20	160	6.3	346 1	SYW_CHLMU
21	159.5	6.3	337 1	SYW_TREPA
22	155	6.1	1714 1	SYEP_DROME
23	153	6.0	1440 1	SYEP_HUMAN
24	152	6.0	344 1	SYW_CHLPR
25	146	5.7	326 1	SYW_HELPY
26	138	5.4	319 1	SYW_METTH
27	137	5.4	528 1	SYW_HUMAN
28	136.5	5.4	528 1	SYW_BOVIN
29	135.5	5.3	328 1	SYW_THEMA
30	133	5.2	328 1	SYW_BACST
31	132.5	5.2	343 1	SYW_MYCLE
32	131	5.1	379 1	SYWM_YEAST
33	130	5.1	326 1	SYW_HELPY

34	129	5.1	401 1	SYWC_SCHPO	O14055 schizosacch
35	128	5.0	334 1	SYW_HAEIN	P43835 haemophilus
36	124.5	4.9	347 1	SYW_MYCE	P47372 mycoplasma
37	122	4.8	350 1	SYWM_CAEPL	P46579 ctenorhadi
38	118.5	4.7	343 1	SYW_CLOAB	O97166 clostridium
39	116.5	4.6	330 1	SYW_BACSO	P21656 bacillus su
40	112.5	4.4	335 1	SYW_BUCAT	P57602 buchara ap
41	111.5	4.4	336 1	SYW_MYCAT	O53386 mycobacteri
42	111.5	4.4	1067 1	IMB4_SCHPO	O60100 schizosacch
43	110	4.3	306 1	SYW_METUA	O57834 methanococc
44	107.5	4.2	1088 1	RRPO_ROTBR	P17468 bovine rota
45	107	4.2	330 1	SYW_RICPR	O92d76 rickettsia

## ALIGNMENTS

RESULT 1	ID	SYW_HUMAN	STANDARD:	PRT:	471 AA.
AC	P23381	P78535: Q9UDL3:			
DT	01-NOV-1991	(Rel. 20, Created)			
DT	01-DEC-1992	(Rel. 24, Last sequence update)			
DT	16-OCT-2001	(Rel. 40, Last annotation update)			
DE	Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA ligase)				
DE	(TrpRS) (IFP53) (hmRS).				
CN	WARS.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=92105071: PubMed=1761529;				
RA	Rubin B.Y., Anderson S.L., Xing L., Powell R.J., Tate W.P.;				
RT	"Interferon Induces tryptophanyl-tRNA synthetase expression in human				
RT	fibroblasts.";				
RL	J. Biol. Chem. 266:24245-24248(1991).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=92107982: PubMed=1763065;				
RA	Flecker J., Rasmussen H.H., Justesen J.;				
RT	"Human interferon gamma potentially induces the synthesis of a 55-kDa				
RT	protein (gamma 2) highly homologous to rabbit peptide chain release				
RT	factor and bovine tryptophanyl-tRNA synthetase.";				
RL	Proc. Natl. Acad. Sci. U.S.A. 88:11520-11524(1991).				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=92110558: PubMed=1765274;				
RA	Frolova L.Y., Sudomolina M.A., Grigorjeva A.Y., Zinovleva O.L.,				
RT	Kisselev L.L.;				
RT	"Cloning and nucleotide sequence of the structural gene encoding for				
RT	human tryptophanyl-tRNA synthetase.";				
RL	Gene 109:291-296(1991).				
RN	[4]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=92164636: PubMed=1537332;				
RA	Buwaitt U., Plohr T., Boettger E.C.;				
RT	"Molecular cloning and characterization of an interferon induced				
RT	human cDNA with sequence homology to a mammalian peptide chain				
RT	release factor.";				
RL	EMBO J. 11:489-496(1992).				
RN	[5]				
RP	SEQUENCE OF 1-13 FROM N.A.				
RX	MEDLINE=96319994: PubMed=8724762;				
RA	Sokolova I.V., Narovilianskii A.N., Amchenkova A.M., Turpaev K.T.;				
RT	"Alternative splicing of 5'-terminal exons of the human tryptophanyl-				
RT	tRNA synthetase gene.";				
RL	Mol. Biol. (Mosk) 30:319-329(1996).				
RN	[6]				
RP	SEQUENCE OF 1-141 AND 182-471 FROM N.A.				
RC	TISSUE=Spem;				
RX	MEDLINE=93292992: PubMed=7685728;				

Db 342 IAMEKLERFLAEHOSRLEKAKTIAMKLVPPR 373

Search completed: October 24, 2002, 12:53:51  
Job time : 18.2689 secs

Db 331 ECKNNTAEIRKFFEEELSVKREK 353

RESULT 13

Trypophan--tRNA ligase (EC 6.1.1.2) [imported] - Encephalitozoon cuniculi (fragment)

C:Species: Encephalitozoon cuniculi

C:Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 02-Sep-2000

C:Accession: T43806

R:Peptrellidae, E.; Broussolle, V.; Peyret, P.; Metenier, G.; Gouy, M.; Vivares, C.P.

Mol. Biol. Evol. 15, 683-689, 1998

A:Title: Microsporidia, amitochondrial protists, possess a 70-kDa heat shock protein gene

A:Reference number: 22693; MUID:98277683

A:Accession: T43806

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-134 <PEV>

A:Cross-references: EMBL:AJ012470; PIDs:CA10034.1

C:Superfamily: mammalian tryptophan--tRNA ligase; amino acid--tRNA ligase repeat homolog

C:Keywords: ligase

Query Match 15.2%; Score 386; DB 2; Length 134;

Best Local Similarity 51.5%; Pred. No. 4.8e-23;

Matches 68; Conservative 25; Mismatches 35; Indels 4; Gaps 1;

QY 82 EDFVPTVOTSSAK---GIDYDKLIVFGSSKIDKELINRIERATGQRPHHLRGI 137

Db 3 EQRITPMVEVYSDVEVALDIDYKIIQNGCEKFNOLADRLKSLGKRAHYFRRGIV 62

QY 138 FSHRDMNOVDAYENKKPFYLYTGRGSSSEAMVHGHLIPFTFKWLODVNPVLVIO 197

Db 63 FAHDFNLLDEIANNRPFYLYTGRGSSKTMHIGHTIPFLCKYMODAFKRLRIQT 122

QY 198 DEKYLWKDITLD 209

Db 123 DEKFLMSKRL 134

RESULT 14

E69461

tryptophan]-tRNA synthetase (trps) homolog - Archaeoglobus fulgidus

C:Species: Archaeoglobus fulgidus

C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 21-Jul-2000

C:Accession: E69461

R:Kleink, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson

.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirlness, E.F.

Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.

Nature 390, 364-370, 1997

A:Authors: Uterback, T.; Cotton, M.D.; Spriggs, T.; Artlisch, P.; Kaine, B.P.; Sykes, S.

Smith, H.O.; Woese, C.R.; Venter, J.C.

A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archae

A:Reference number: A69250; MUID:98049343

A:Accession: E69461

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-420 <KLE>

A:Cross-references: GB:AE000986; GB:AE000782; NID:92689309; PIDs:AAH8954.1; PID:9264885

C:Superfamily: mammalian tryptophan--tRNA ligase; amino acid--tRNA ligase repeat homolog

Query Match 14.5%; Score 370.5; DB 2; Length 420;

Best Local Similarity 27.3%; Pred. No. 4.2e-21;

Matches 118; Conservative 60; Mismatches 166; Indels 89; Gaps 12;

QY 85 VDPPTVOTSSAKGIDYDKLIVFGSSSKIDKELINRIERATGQRPHHLRGI 144

Db 3 VTPNEVEGV---IDYSLIEFGMOFP-SEVLEPID-----NPHILMRGAIIGHRDY 52

QY 145 QVLDAVENKKPFYLYTGRGSSSEAMVHGHLIPFTFKWLODVNPVLVIO 204

Db 53 RITAMOKKEKEMVAVSGMPSG-LPHFGHKMTMDIYVHOSAGKAFV-ATADMAHSVR 110

QY 205 DLTLDQAYGDAVENAKDIACGFDINKTFISDLDYMGSSGFYKNV-KIOKVTFTNOV 263

Db 111 GLSWEKTELMGLYIKSIILALREDVAIYFQS-----KSHVYKDLAFELSAEVNSEL 164

QY 264 KGICGFTDSDICIGKISFPALAAPSFNSPFOIRDRDIOCLIPCAIDOPYFRMRDV 323

Db 165 RAIFGFNSDTSLAKMFYTAIOADTL---HPQLSDFGGPRKVVVPVAGADDPHMKLRDL 221

QY 324 APR1----- 327

Db 222 AARISIFSEFEVEGGVRRSKGAELSLDLDFDKKIYEHNMDIGEAEIERAVRKI 281

QY 328 -----GYRPPALLHSTFPALOGAQTMSASDPNSSIFLTPTAKQIKTVKNHATSG 380

Db 282 EVEIGCFAPFIPPSSTYHRTFTGLTG--KMSSSKPEYSISLDPPEEGAKKVMK-AFTGG 338

QY 381 RDTIEHROFGNDDVVSFMYLTFPFLD--DDKLEQIRKDDVTSGAMLTGELKRALIEYLO 439

Db 339 RATAEQRRLGGEPRDRCVVFELYSFHLSDSEELNOLAECEBREGRLCGKCKKMAELVK 398

QY 440 PLIAEHQARKEV 452

Db 399 SFLKEHQKMEAV 411

RESULT 15

D72477

probable tryptophanyl-tRNA synthetase APE2461 - Aeropyrum pernix (strain K1)

C:Species: Aeropyrum pernix

C:Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 20-Jun-2000

C:Accession: D72477

R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Ta

awa, H.; Takamizawa, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.

DNA Res. 6, 83-101, 1999

A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aero

A:Reference number: A72450; MUID:99310339

A:Accession: D72477

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-374 <KAM>

A:Cross-references: DBJ:AP000064; NID:95105945; PIDs:BAA81476.1; PID:95106165

A:Experimental source: strain K1

C:Genetics:

A:Gene: APE2461

C:Superfamily: mammalian tryptophan--tRNA ligase; amino acid--tRNA ligase repeat homo

Query Match 13.9%; Score 353; DB 2; Length 374;

Best Local Similarity 30.4%; Pred. No. 8.2e-20;

Matches 119; Conservative 70; Mismatches 165; Indels 38; Gaps 17;

QY 85 VDPPTVOTSSAKGIDYDKLIVFGSSSKIDKELINRIERATGQRPHHLRGI 144

Db 8 LDPW---GAVEIKDYDRLRTFGIRPSEVL--PLRKAGMEPSPLMRGIITGHRDP 61

QY 145 QVLDAVENKKPFYLYTGRGSSSEAMVHGHLIPFTFKWLO-DVNPVLVIO 203

Db 62 KILEAKARGERVAVALTGMPSG-FHFGHKLTVDOLLYLQNGKRV--FVAIADEAFVAV 118

QY 204 KDLLDQAYGDAVEN-AKDIACGFDINKT-FISDLDYMGSSGFYKNV-KIOKVTFTN 261

Db 119 RIRIREEVRIAYEVIYANMTALGLDPKDTIEYQ---TKRGIPYRLLQJESGKYTAA 174

QY 262 QVKGIFG-FETSDSDICIGKISFPALAAPSFNSPFOIRDRDIOCLIPCAIDOPYFRMT 320

Db 175 EMEATYGEIPLPAKMMASLT-----QAQDLNVQDLEYGYR--HVVVPVAGADDPHRLRT 227

QY 321 RDVAPR---IGYRPPALLHSTFPALOGAQTMSASDPNSSIFLTPTAKQIKTVKNKA 376

Db 228 RDLADRMAGVVELRPPASTYHKLQPLGD--RKMSSSRPDSTIFLTDPPEVARRKLR-A 284

QY 377 FSGGRDITIEHROFGKCDV-DVSFMYLTFLEDDKDLQIRKQYTS---GAMLTGELK 432

Db 285 LTGGRATAEQRRLGCGVEVSVYHMDLYHLMPPDDGEVKNH--YTSGRILKILGCECKQ 341

QY 433 ALIEVLOPLIAEHQARKEVTDIEVKEMPR 464

**Query Match** 21.0% Score 535.5; DB 2; Length 380;  
Best Local Similarity 35.2%; Pred. No. 4.6e-34;  
**Matches** 142; Conservative 78; Mismatches 142; Indels 41; Gaps 17;

Oy AAEEDVDPMFVOTSSAKGIDYDKLIVRFGSSKIDKELINIRIERTATGORPHHRLRGIFES 139  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db ADGNVTYTYAAVESD---LDYEKLARFGADELTDDQARRP-----DHPLVNKGLEYA 53  
|||:::|||||:  
Oy HRDMNQVLDAVENKKPFYLITGRGPSEAMHVGHILIPETFTKWLDVEVNEVLDTOMTDE 199  
|||:::|||||:  
Db GRDVDDFLTAGBQS----LYTVGVPSG--PHMIGHAMVFYFARRLODERGARAVYPVLSDDE 108  
|||:::|||||:  
Oy KYLMKDLLDAQYGDAVE-NAKDIACGFDPINKPIFI---SDDLVM-GMSGCFKNKYVKI 254  
|||:::|||||:  
Db KWFMEQQTPAET-GDYLANLRDLVAWGEDELFRIVDTADADVLPALAFAGADV--- 164  
|||:::|||||:  
Oy QKHVPENOVKGIFFGTDSDCIGKSFPALIOAPRSNSFPOLFNRNPDIOCLIPAIODD 314  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db RHATLOWNYC-----EPDNVGQAFYPAVOITHLL--LPOLVHG--EHETLVPIAYVOD 213  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Oy PYFMRTBDVABRICYP--KPALLNSTFFPALOGAQTKMSASDPNSSIFIPTTAOKIKRV 372  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db PHVRSRBVAAKARKRPVKGPGALLMQFLPSLAG-PGKMSSS-AAGSIKLTSPDVRREKV 271  
|||:::|||||:  
Oy NKHAFSGGRDITEEHROFGNCDDVVSVFWYLFTELEDD-KLEQIRKDYTSGAMLTELK 431  
|||:::|||||:  
Db RTHATYGSRASVEEHRRAKGVAPEDEVPOYLSAFEPPDAELARIERYRAGDLSSELK 331  
|||:::|||||:  
Oy KALLEVOLPLIAHQRRKEVTDIEVKEFMTRPKISFPQKLA 474  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db DLADRITFEFLAHQRRAALGD-VTEALDFRLTDERORA 372  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :

**RESULT 11**  
F64476  
tryptophan--tRNA ligase (EC 6.1.1.2) - Methanococcus jannaschii  
N:Alternate names: tryptophanyl-tRNA synthetase  
C:Species: Methanococcus jannaschii  
C:Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 21-Jul-2000  
C:Accession: F64476  
R:Built C.J.: White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake,  
; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstein, K.G.; Metrick, J.M.; Giodek, A.  
rson, J.D.; Sadov, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.  
tsom, J.D., 1996  
A:Authors: Kalne, B.P.; Borodovsky, M.; Kleen, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C  
A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii  
A:Reference number: A64300; MUID:96373999  
A:Accession: F64476  
A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-370 <BU>  
C:Genetics:  
A:Map position: FOR1375865-1376997  
A:Start codon: GTG  
C:Superfamily: mammalian tryptophan--tRNA ligase; amino acid--tRNA ligase repeat homolog  
C:Keywords: aminocacyl-tRNA synthetase; ligase; protein biosynthesis

**Query Match** 16.1% Score 409.5; DB 2; Length 370;  
Best Local Similarity 30.9%; Pred. No. 5.1e-24;  
**Matches** 119; Conservative 72; Mismatches 153; Indels 41; Gaps 16;

Oy PMTVOTSAGKGIDYDKLIVRFGSSKIDKELINIRIERTATGORPHHRLRGIFESHMDMNQV 146  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db PW-ETPRV-IDYKTMEQFCVKRIVDGLDAKE-----HHFFRNIIIGHRDEFERI 57  
|||:::|||||:  
Oy LDVAENKKPFYLITGRGPSEAMHVGHILIPETFTKWLD--DVFNVLPIQMTDEKYLEM 203  
|||:::|||||:  
Db VDAIKNKKEFAVASGMPSGK-MHFCHKMVVDLKFQYKYNIDINIPI---ADLEAWYA 112  
|||:::|||||:  
Oy KDULTIQAYGDAV-ENADIITACGVDINKTFIFSOLDYMGWSSGCTKANVKA-QKHVTFN 261  
|||:::|||||:  
Db RNMSSETTKELALNIYINYIALGIDPERKINVYLOSCKÖKY---KDLAILSLSKRTWS 167  
|||:::|||||:

QY 262 QYKQIFGFTDSCITGKISFPALQAAAPSFNSFPQIFRORT--DIQCLIPCALDDPVRFM 319  
 Db 168 EKKAIYGGKGTETNTHVAPRIQVADIL--HPQDENLSEPPKRVVPVAVGIDDPHRL 224  
 QY 320 TRDVAIPR--IGYKPKALHSTFPALOGAQTAKMSASPNSSIFLTPDAKQIKRVNHA 376  
 Db 225 TRDIANRAKKEKFPITPPSSYTHFMGTLGG--KMSSSRPETALFLTLDKDEKIVKKIPS-A 281  
 QY 377 FSGGDTIEEHRQFG--NCDVDSFMVLTFFLEDKLEQIRKDYSGAMLTGELKA 433  
 Db 282 KKGCGRETFLEEHKKYGVGVEECVYELFLY--HLIUDKELAEIYQKCRSGELTGCCKKM 339  
 QY 434 LIEVQPLIAEHQARKEVTDIYK 458  
 Db 340 AYERVEFLKDLKEKREQAKELAVK 364

RESULT 12  
 E69131  
 tryptophan--tRNA ligase (EC 6.1.1.2) - Methanobacterium thermoautotrophicum (strain D  
 N:Alternate names: tryptophanyl-tRNA synthetase  
 C:Species: Methanobacterium thermoautotrophicum  
 C:date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 22-Oct-1999  
 C:Accession: E69131  
 R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T  
 ; Qiu, D.; Spadafora, R.; Vitale, R.; Wang, Y.; Mierzbowski, J.; Gibson, R.; Jivan, I.  
 K.; S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Neelling, J.; Reeve, J.N.  
 J. Bacteriol. 179, 7135-7155, 1997  
 A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H:  
 A:Reference number: A69000; MUID:98037514  
 A:Accession: E69131  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-364 <MTH>  
 A:Cross-references: GB:A6000812; GB:A6000666; NID:92621298; PIDN:AA84757.1; PID:g2622  
 A:Experimental source: strain Delta H  
 C:Genetics:  
 A:Gene: MTH251  
 A:Start codon: TTG  
 C:Superfamily: mammalian tryptophan--tRNA ligase; amino acid--tRNA ligase repeat homo  
 C:Keywords: aminocacyl-tRNA synthetase; ligase; protein biosynthesis

Query Match 15.6%; Score 397.5; DB 2; Length 364;  
 Best local similarity 27.9%; Pred. No. 2.6e-23;  
 Matches 107; Conservative 76; Mismatches 153; Indels 47; Gaps 14;

QY 85 VDPWTVQTSSAKGIDYDKLIVFGSSKIDKELINRTERATGQRPHFLRGIFFSHRDMN 144  
 Db 2 IDPW---GSAK-LEYPDLEINFGVRPF-SEVLDEV-----PEPSWILMRRGIIIGHRYE 50  
 QY 145 QYLDVAENKKPRYLITGKSPSEAMHVGHLIPETITKALQDYFNVPLVITQTTDDEKYLK 204  
 Db 51 RLISAMKKGEDPAVVTGGMPPSGR-MHIGKMLVQDLRW-YRMCAGELRPIADMEAVSAR 108  
 QY 205 DLTLOAGDAVEN-AKDIACGFDINK-----TFLPSDDYDGMSGFGKVVVK 253  
 Db 109 GVDPEDSRIALIEEYIAGYIALGDLDEKDNHVIYLSQENLVEDIAY----- 156  
 QY 254 IQKAVTFNQVKGIFGFTDSDCTGKISFPALQAAAPSFNSFPQIFRRTDIOCLIPCALDQ 313  
 Db 157 LAGKVNENELRAIYGTGSTSMAHMYARILIQSDILHPQDLDELGRP---PVIVPVGPDQ 213  
 QY 314 DYEFMTRDVAPRI---GYKPKALHSTFPALOGAQTAKMSASPNSSIFLTPDAKQIK 369  
 Db 214 DPHILTRDIARAFRDYRGFILLPSSTYHRFMGGLGG--KMSNMKPSAIFLSDTPEAE 271  
 QY 370 TVVNHAFSGSGDITIEEHRQFGNCVDVDSFMVLTFFLE-DDDKLEQIRKDYTSGAMLTG 428  
 Db 272 AKI-RNAATGCGETLKEQRELQGVPEECIYETTLIYHSGSDSRULEIYBSCRNTLMCG 330  
 QY 429 ELKKALIEVLQPLIAEHQARKE 451

Db 119 DDEKYMNRDEFTLDOTSRMAYNIIIDIIAVGFNPDKTEIFODPEXI---RNMYPITVAKIA 175  
Oy 256 KHVTFNOKGIFGFTDSDICIGTISFPALQAAFSNSFPQIRDRDFTDIOCLIPCAIDDDP 315  
Db 176 KKLATSEVRAFEGLDASSNIGTIFPALQIAPT-----MEKR---RCLIPALDIDDP 225  
Oy 316 YFRMRDVAIRIGYKPKPALHSTFFPALQGAOTKMSASDPNSISILFTJAKOIKRVKMH 375  
Db 226 YWRLOORDAESIGYKKAQIHKFLPRLTGPBGKSSNSPETAIRIYLVDPDKTVERKIKY 285  
Oy 376 AFSGGRDTIEEHROFGNCDVDVSMYLTFFLEDDD-KLEQIRKDYTGAMLTGELKKAL 434  
Db 286 AFSGQGPTEILCHRRYKGGMEIDVFPQWLYFFFEEDDNRIKTEIEEYRSGKMLTGELKQIL 345  
Oy 435 IEVLQPLAEHQARKKEVTDDEIVKKEFMTPRKLS 467  
Db 346 IDKLNLNLEBHR-RRREKAEKLHVFKYDGKLA 377

## RESULT 8

C75020  
tryptophanyl-tRNA synthetase (trps) PAB111 - Pyrococcus abyssi (strain Orsay)  
C:Species: Pyrococcus abyssi  
C:Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 20-Jun-2000  
C:Accession: C75020  
R:anonymous, Genoscope  
Submitted to the EMBL Data Library, July 1999  
A:Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome stru  
A:Reference number: A75001  
A:Accession: C75020  
A:Status: Preliminary  
A:Molecule type: DNA  
A:Residues: 1-385 <KAM>  
A:Cross-references: GB:A2248288; GB:AL096836; NID:g5458960; PIDN:CA850601.1; PID:g545911  
A:Experimental source: Strain Orsay  
C:Genetics:  
A:Gene: trps; PAB111  
C:Superfamily: mammalian tryptophan--tRNA ligase; amino acid--tRNA ligase repeat homolog

Query Match 31.5%; Score 803; DB 2; Length 385;  
Best Local Similarity 45.38; Pred. No. 5.7e-55;  
Matches 173; Conservative 63; Mismatches 120; Indels 26; Gaps 9;

Oy 82 EDF-DPMTVGTSSAKGIDYDKLIYRFGSSKIDKELINRERATGQRPHNLRGIFPSH 140  
Db 3 EDFKTPMEVEGV---VDYKLIHFGTSPLEBELKTAALTSELPLFFRRKRFESH 58  
Oy 141 RDMNOVLDAENKKPFYLYTGRGSSSEAMHVGHLIPFIETKWLDVFNVLVQMTDDEK 200  
Db 59 RQYDKVLQDYEGRGFFLYTGRGSPG-PMHIGHIIPFATKWLQERKFGVNLXIQTDEK 117  
Oy 201 YLMKD-LTLDOAYGDAVENAKDIIACGPDINTFTFSOLDVMGMSGFGYKNVKKQKHVT 259  
Db 118 FLEKFNLFEDTKHAWYENKILDIIVAGFDPDKTFFQNSEF---TKIYEMALIPAKIN 173  
Oy 260 FNOVVGIFGFTDSDICIGTISFPALQAAFSNSFPQIRDRDFTDIOCLIPCAIDDDPYRM 319  
Db 174 FSKMAVAVGFEGSQSKIGMIFFPALQIAPT-----FEKR---RCLIPALDIDDPYRL 223  
Oy 320 TRDVAIRIGYKPKPALHSTFFPALQGAOTKMSASDPNSISILFTJAKOIKRVKMHAFSG 379  
Db 224 QRFDAESIGYKKAQIHKFLPRLTGPBGKSSNSPETAIRIYLVDPDKTVERKIKYFALTG 283  
Oy 380 GQDTIEEHROFGNCDVDVSMYLTFFLEDDDKLEQIRKDY---TSGAMLTGELKKALIE 436  
Db 284 GRPTLEKREKGEKPEKCVFEMLEIFFEEDDK-KLERYYACKNGELTGECCKRYLIS 341  
Oy 437 VLQPLAEHQARKKEVTDDEIVK 458  
Db 342 KIQEFLKEHQRRKKAKEQIEK 363

## RESULT 9

G71206  
tryptophan--tRNA ligase (EC 6.1.1.2) - Pyrococcus horikoshii  
C:Species: Pyrococcus horikoshii  
C:Date: 14-Aug-1998 #sequence\_revision 14-Aug-1998 #text\_change 21-Jul-2000  
C:Accession: G71206  
R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Halkawa, Y.; Hino, Y.; Yamamoto, S.; Se  
M.; Ohkuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kishida, N.; Ogu  
DNA Res. 5: 55-76, 1998  
A:Title: Complete sequence and gene organization of the genome of a hyper-thermophil  
A:Reference number: A71000; MUID:98344137  
A:Accession: G71206  
A:Status: Preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-301 <KAM>  
A:Cross-references: GB:AF000007; NID:g3236134; PIDN:BA831046.1; PID:g3258363  
A:Experimental source: Strain OT3  
A:Note: this accession replaces an interim accession for a sequence replaced by Genba  
C:Genetics:  
A:Gene: PH1921  
C:Superfamily: yeast tyrosine--tRNA ligase  
C:Keywords: aminocyl-tRNA synthetase; ligase; protein biosynthesis

Query Match 24.68; Score 626.5; DB 2; Length 301;  
Best Local Similarity 44.28; Pred. No. 2.5e-41;  
Matches 134; Conservative 56; Mismatches 92; Indels 21; Gaps 7;

Oy 169 MHVGLIPFIETKWLDVFNVLVQMTDDEKLYMKD-LTLDOAYGDAVENAKDIIACG 227  
Db 1 MHIGHIIPFATKWLQERKFGVNLXIQTDEKLEKELTDDTKRMAYNDIILDIIVAGF 60  
Oy 228 DINKTIFSDLDYMGMSGFGYKNVKKQKHVTFNQKIGFTDSDICIGTISFPALQAA 287  
Db 61 DPKDTFIQNSEF-----TKIYEMALIPAKINFSMAKAVFETBSKIGMIFFPALQIAP 116  
Oy 288 SFSNSFPQIRDRDFTDIOCLIPCAIDDDPYFRMTRDVAIRIGYKPKPALHSTFFPALQGAQ 347  
Db 117 TF-----FERK---RCLIPALDIDDPYWRLOORDFAESLGYKKAHLSKFSVSLTSL 166  
Oy 348 TKMSASDPNSISILFTJAKOIKRVKMHAFSGGRDTIEEHROFGNCDVDVSMYLTFFEL 407  
Db 167 GKMSASKRETAIVLTDSPEDVEKRWKFTLTGGRPTLEQKRGGEPEKCVFEMLEIF 226  
Oy 408 EDDDKLEQIRKDY---TSGAMLTGELKKALIEVLQPLAEHQARKKEVTDDEIVK 464  
Db 227 EEDDK-KLERYYACKNGELTGECCKRYLISKIQEFLKEHQRRKK-AEKLVEKFKYTG 283  
Oy 465 KLS 467  
Db 284 KLA 286

## RESULT 10

G84373  
tryptophanyl-tRNA synthetase [imported] - Halobacterium sp. NRC-1  
C:Species: Halobacterium sp. NRC-1  
C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 16-Feb-2001  
C:Accession: G84373  
R:Ng, W.V.; Kennedy, S.P.; Mahatras, G.G.; Bergquist, B.; Pan, M.; Shukla, H.D.; Lasky  
; Lelthausen, B.; Keller, K.; Cruz, R.; Danson, M.T.; Hough, D.W.; Maddocks, D.G.; Ja  
Jung, K.H.; Alam, M.; Freitas, T.  
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000  
A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ehardt, H.; Lowe, T.M.;  
A:Title: Genome sequence of Halobacterium species NRC-1.  
A:Reference number: A84160; MUID:20504483  
A:Accession: G84373  
A:Status: Preliminary  
A:Molecule type: DNA  
A:Residues: 1-380 <STO>  
A:Cross-references: GB:AE004437; NID:g10581646; PIDN:AA620355.1; GSPDB:GN00138  
C:Genetics:  
A:Gene: trps2  
C:Superfamily: mammalian tryptophan--tRNA ligase; amino acid--tRNA ligase repeat homo







A:Residues: 1-212,'GD',215-471 <PRO2>  
 A:Cross-references: GB:M61715; NID:g340367; PIDN:AAA61298.1; PID:g340368  
 A:Experimental source: fibroblast  
 C:Genetics:  
 A:Gene: GDB:WARS; IFP53  
 A:Cross-references: GDB:119632; OMIM:191050  
 A:Map position: 14q23-14q31  
 A:Introns: 33/3; 105/1; 141/2; 181/2; 242/2; 276/1; 313/3; 371/3; 418/3  
 C:Superfamily: mammalian tryptophan--tRNA ligase; amino acid--tRNA ligase repeat homolog  
 C:Keywords: aminocyl-tRNA synthetase; ATP; ligase; protein biosynthesis  
 F:19-64/Domain: amino acid--tRNA ligase repeat homology <ATL>

Query Match 96.3%; Score 2454; DB 1; Length 471;  
 Best Local Similarity 99.6%; Pred. No. 6.7e-184;  
 Matches 469; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 MNSPSPASILEFNSIATQGEIVRSLSKAGNASKDEIDSAVKMLVSLKMSYKAAAGEDYKA 60  
 |||||||  
 DB 1 MNSPSPASILEFNSIATQGEIVRSLSKAGNASKDEIDSAVKMLVSLKMSYKAAAGEDYKA 60

OY 61 DCPENPAPTSNMGPDATAEEDFVDPMTVOTSSAKGIDYDKLIYRFGSSKIDKELINRI 120  
 |||||||  
 DB 61 DCPENPAPTSNMGPDATAEEDFVDPMTVOTSSAKGIDYDKLIYRFGSSKIDKELINRI 120

OY 121 ERATGQRPHHFLRGIFFSHRDMNOVLDAVENKKRPYLYTGRPSSEAMHVGHLIPEFT 180  
 |||||||  
 DB 121 ERATGQRPHHFLRGIFFSHRDMNOVLDAVENKKRPYLYTGRPSSEAMHVGHLIPEFT 180

OY 181 KWLQDVFNPLVYQMTDDEKYLKDLTLDOAYGAVENAKDITACGPDINKTFIESDLY 240  
 |||||||  
 DB 181 KWLQDVFNPLVYQMTDDEKYLKDLTLDOAYGAVENAKDITACGPDINKTFIESDLY 240

OY 241 MMSGFFKNNVKIQKHVFNQVKGIFGFTSDSCIGKISFPALQAPSPNSFPQIFRDR 300  
 |||||||  
 DB 241 MMSGFFKNNVKIQKHVFNQVKGIFGFTSDSCIGKISFPALQAPSPNSFPQIFRDR 300

OY 301 TDIOCLIPCAIDDDPYFRMTRDVAPRIGYKPKALLHSTFFPALQAGQTKMSADPNSSIF 360  
 |||||||  
 DB 301 TDIOCLIPCAIDDDPYFRMTRDVAPRIGYKPKALLHSTFFPALQAGQTKMSADPNSSIF 360

OY 361 LTDPAKQIKTKVKNKHAFFSGGRDTIEHRQFGNCDDVDFSMILTFLEDDDKLEQIRKDY 420  
 |||||||  
 DB 361 LTDPAKQIKTKVKNKHAFFSGGRDTIEHRQFGNCDDVDFSMILTFLEDDDKLEQIRKDY 420

OY 421 TSGAMLTGELKALIEVLOPLIAEHQARRKEVTEIVKEKEMTPRKLSDFQ 471  
 |||||||  
 DB 421 TSGAMLTGELKALIEVLOPLIAEHQARRKEVTEIVKEKEMTPRKLSDFQ 471

RESULT 2  
 YWBO  
 N:Alterphan--tRNA ligase (EC 6.1.1.2) [validated] - bovine  
 C:Species: Bos primigenius taurus (cattle)  
 C:Date: 30-Sep-1992 #sequence\_revision 30-Sep-1992 #text\_change 21-Jul-2000  
 C:Accession: A40279; JN0354; S10460; S14540  
 R:Garret, M.; Pajot, B.; Treguelet, V.; Labouesse, J.; Merle, M.; Gandar, J.C.; Benedetti  
 Biochem J. 30, 7809-7817, 1991  
 A:Title: A mammalian tryptophanyl-tRNA synthetase shows little homology to prokaryotic  
 A:Reference number: A40279; MUID:91329348  
 A:Accession: A40279  
 A:Molecule type: mRNA  
 A:Residues: 1-475 <GAR>  
 A:Cross-references: GB:M74074; EMBL:X53918; NID:g163798; PIDN:AAA30799.1; PID:g163799  
 A:Experimental source: pancreas  
 A:Note: the authors translated the codon CTG for residue 347 as Ala and CAG for residue  
 R:Zatagova, T.A.; Kovaleva, G.K.; Favorova, O.O.; Aevina, H.B.; Telezhinskaya, I.H.  
 Bioorg. Khim. 15, 1307-1311, 1989  
 A:Title: Amino acid sequence of several peptides of tryptophanyl-tRNA synthetase from ca  
 A:Accession: JN0354; MUID:90211408  
 A:Molecule type: protein

A:Residues: 112-124;282-287,'N',288,'F',289-292,'Q',293-294,'IR',336-353;423-441,443-  
 A:Experimental source: liver  
 A:Note: this paper is in Russian  
 C:Superfamily: mammalian tryptophan--tRNA ligase; amino acid--tRNA ligase repeat homo  
 C:Keywords: aminocyl-tRNA synthetase; ATP; ligase; protein biosynthesis  
 F:24-69/Domain: amino acid--tRNA ligase repeat homology <ATL>

Query Match 89.8%; Score 2286.5; DB 1; Length 475;  
 Best Local Similarity 92.6%; Pred. No. 8.4e-171;  
 Matches 438; Conservative 12; Mismatches 20; Indels 3; Gaps 2;

OY 1 MNSPSPASILEFNSIATQGEIVRSLSKAGNASKDEIDSAVKMLVSLKMSYKAAAGEDY 58  
 |||||||  
 DB 4 MNSGQGGCGSPLEFHSIAOGEIVRDLKARNAADEIDSAVKMLLSKTYSKATGEDY 63

OY 59 KADCPGNPAPTSNMGPDATAEEDFVDPMTVOTSSAKGIDYDKLIYRFGSSKIDKELIN 118  
 |||||||  
 DB 64 KYDCPPGPAPSPGGLDTEADEDFVDPMTVOTSSAGIDYDKLIYRFGSSKIDKELVN 123

OY 119 RIERATGQRPHHFLRGIFFSHRDMNOVLDAVENKKRPYLYTGRPSSEAMHVGHLIPEFT 178  
 |||||||  
 DB 124 RIERATGQRPHHFLRGIFFSHRDMNOVLDAVENKKRPYLYTGRPSSEAMHVGHLIPEFT 183

OY 179 FTKWLQDVFNPLVYQMTDDEKYLKDLTLDOAYGAVENAKDITACGPDINKTFIESDL 238  
 |||||||  
 DB 184 FTKWLQDVFNPLVYQMTDDEKYLKDLTLDOAYGAVENAKDITACGPDINKTFIESDL 242

OY 239 DYMMSGFFKNNVKIQKHVFNQVKGIFGFTSDSCIGKISFPALQAPSPNSFPQIFRDR 298  
 |||||||  
 DB 243 DYMMSGFFKNNVKIQKHVFNQVKGIFGFTSDSCIGKISFPALQAPSPNSFPQIFRDR 302

OY 299 DTDIOCLIPCAIDDDPYFRMTRDVAPRIGYKPKALLHSTFFPALQAGQTKMSADPNSS 358  
 |||||||  
 DB 303 DTDIOCLIPCAIDDDPYFRMTRDVAPRIGYKPKALLHSTFFPALQAGQTKMSADPNSS 362

OY 359 IFLDTAOKIKTKVKNKHAFFSGGRDTIEHRQFGNCDDVDFSMILTFLEDDDKLEQIRK 418  
 |||||||  
 DB 363 IFLDTAOKIKTKVKNKHAFFSGGRDTIEHRQFGNCDDVDFSMILTFLEDDDKLEQIRK 422

OY 419 DYTSGAMLTGELKALIEVLOPLIAEHQARRKEVTEIVKEKEMTPRKLSDFQ 471  
 |||||||  
 DB 423 DYTSGAMLTGELKALIEVLOPLIAEHQARRKEVTEIVKEKEMTPRKLSDFQ 475

RESULT 3  
 S50053  
 tryptophan--tRNA ligase (EC 6.1.1.2) alpha-2 chain - mouse  
 C:Species: Mus musculus (house mouse)  
 C:Date: 07-May-1995 #sequence\_revision 10-Nov-1995 #text\_change 16-Jul-1999  
 C:Accession: S50053; S50052; I49391; S31461; S31462  
 R:Pajot, B.; Sarger, C.; Bonnet, J.; Garret, M.  
 J. Mol. Biol. 242, 599-603, 1994  
 A:Title: An alternative splicing modifies the C-terminal end of tryptophanyl-tRNA syn  
 A:Reference number: S50052; MUID:95018226  
 A:Accession: S50053  
 A:Molecule type: mRNA  
 A:Residues: 1-481 <PAV>  
 A:Cross-references: EMBL:X69657; NID:g55437; PIDN:CAA49348.1; PID:g55438  
 A:Genetics: LSP  
 A:Note: Intron position was determined by sequencing of genomic DNA  
 A:Accession: S50052  
 A:Molecule type: mRNA  
 A:Residues: 1-475 <PAV>  
 A:Cross-references: EMBL:X69656; NID:g55435; PIDN:CAA49347.1; PID:g55436  
 A:Genetics: SSP  
 R:Kisselev, L.L.  
 Biochimie 75, 1027-1039, 1993  
 A:Title: Mammalian tryptophanyl-tRNA synthetases.  
 A:Reference number: I49391; MUID:94257729  
 A:Accession: I49391  
 A:Status: preliminary; translated from GR/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-481 <RES>

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

# OM protein - protein search, using sw model

Run on: October 24, 2002, 12:49:45 : Search time 16.2689 Seconds  
(without alignments)  
2858.658 Million cell updates/sec

Title: US-09-813-718-10  
Perfect score: 2547  
Sequence: 1 MPNSEPALSLELFNSIATQG.....KLSFDQKLAALHHHHHH 484

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :  
1: PIR\_71:\*  
2: PIR1:\*  
3: PIR3:\*  
4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2454	96.3	471	1 A41706	tryptophan--trna 1
2	2286.5	89.8	475	1 YWBO	tryptophan--trna 1
3	2208.5	86.7	481	1 S50053	tryptophan--trna 1
4	2177	85.5	475	1 YWRBPR	tryptophan--trna 1
5	1210	47.5	395	2 S58157	hypothetical prote
6	1163	45.7	432	2 S51901	tryptophan--trna 1
7	907	35.6	386	2 C90190	tryptophan--trna 1
8	803	31.5	385	2 C75020	tryptophanyl--trna
9	626.5	24.6	301	2 G71206	tryptophan--trna 1
10	535.5	21.0	380	2 G84373	tryptophanyl--trna
11	409.5	16.1	370	2 F64476	tryptophan--trna 1
12	397.5	15.6	364	2 E69131	tryptophan--trna 1
13	386	15.2	134	2 T43806	tryptophan--trna 1
14	370.5	14.5	420	2 E69451	tryptophanyl--trna
15	353	13.9	374	2 D72477	probable tryptopha
16	274.5	10.8	513	2 F84371	tryptophanyl--trna
17	192	7.5	323	2 H69346	tyrosyl--trna synth
18	188	7.4	364	2 E72512	probable tyrosyl-t
19	185.5	7.3	341	2 D95260	tryptophanyl--trna
20	185.5	7.3	341	2 G98125	tryptophan--trna 1
21	174.5	6.9	341	2 B86633	tryptophan--trna 1
22	172	6.8	394	2 A45999	tyrosine--trna 11g
23	170.5	6.7	366	2 S75410	tyrosine--trna 11g
24	169	6.6	395	2 H70385	tryptophan--trna 1
25	168.5	6.6	895	2 A86410	protein F3M18.22 l
26	167	6.6	351	2 E75438	tryptophanyl--trna
27	166.5	6.5	346	2 B71496	tryptophan--trna 1
28	163.5	6.4	353	2 E70100	tryptophan--trna 1
29	163	6.4	460	2 C84750	probable tyrosyl-t

30	160	6.3	346	2 C81654	tryptophanyl--trna
31	159.5	6.3	337	2 F71300	tryptophan--trna 1
32	158.5	6.2	375	2 B75072	tyrosyl--trna synth
33	155	6.1	1714	1 S18644	multifunctional am
34	154.5	6.1	102	2 T44994	probable tryptopha
35	153	6.0	1440	1 SYHUQT	multifunctional am
36	152	6.0	344	2 H86590	tryptophanyl trna
37	152	6.0	344	2 C72034	tryptophan--trna 1
38	149.5	5.9	327	2 C84374	tyrosyl--trna synth
39	149.5	5.9	408	2 T03741	probable tyrosine-
40	146	5.7	339	2 E64676	tryptophan--trna 1
41	146	5.7	375	2 F71093	tyrosine--trna 11g
42	145	5.7	337	2 A11066	tryptophan--trna 1
43	140.5	5.5	1149	2 T27567	hypothetical prote
44	138	5.4	319	2 H69102	tyrosine--trna 11g
45	135.5	5.3	328	2 C72370	tryptophan--trna 1

## ALIGNMENTS

RESULT 1  
A41706  
tryptophan--trna ligase (EC 6.1.1.2) [similarity] - human  
N:Alternate names: Interferon-inducible protein IFP53; peptide-chain release factor h  
C:Species: Homo sapiens (man)  
C>Date: 19-May-2000 #sequence\_revision 19-May-2000 #text\_change 19-May-2000  
C:Accession: A41633; A41706; S19246; JN0676; JH0533; S26287  
R:Fleckner, J.; Rasmussen, H.H.; Justesen, J.  
Proc. Natl. Acad. Sci. U.S.A. 88, 11520-11524, 1991  
A:Title: Human interferon gamma potentially induces the synthesis of a 55-kDa protein (9  
A:Reference number: A41633; MUID:92107982  
A:Accession: A41633  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-471 <RUB>  
A:Cross-references: GB:559892; NID:930820; PIDN:CAA42545.1; PID:930821  
A:Residues: 1-471 <FILE>  
A:Cross-references: GB:559892; NID:930820; PIDN:CAA42545.1; PID:930821  
R:Rubin, B.Y.; Anderson, S.L.; King, L.; Powell, R.J.; Tate, W.P.  
J. Biol. Chem. 266, 24245-24248, 1991  
A:Title: Interferon induces tryptophanyl-trna synthetase expression in human fibrobla  
A:Reference number: A41706; MUID:92105071  
A:Accession: A41706  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-471 <RUB>  
A:Cross-references: GB:M77804; NID:g184656; PIDN:AAA67324.1; PID:g184657  
R:Buwilt, U.; Flom, T.; Boeltger, E.C.  
EMBO J. 11, 489-496, 1992  
A:Title: Molecular cloning and characterization of an interferon induced human cDNA w  
A:Reference number: S19246; MUID:92164636  
A:Accession: S19246  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-423, 'R', 425-471 <BUW>  
A:Cross-references: EMBL:X62570; NID:932708; PIDN:CAA44450.1; PID:932709  
A:Note: 213-Ser and 214-Tyr were also found  
R:Frolova, L.Y.; Grigorjeva, A.Y.; Sudomoina, M.A.; Kisseliev, L.L.  
Gene 126, 237-245, 1993  
A:Title: The human gene encoding tryptophanyl-trna synthetase: Interferon-response el  
A:Reference number: JN0676; MUID:93292992  
A:Accession: JN0676  
A:Molecule type: DNA  
A:Residues: 1-141,182-471 <PROI>  
A:Cross-references: GB:X67918; GB:S62837; NID:g37968; GB:X67919; NID:g37969; GB:X6792  
4; NID:g37974; GB:X67925; GB:S62855; NID:g37975; GB:X67926; GB:S62856; NID:g37976; GB  
A:Note: the authors translated the codon GGC for residue 55 as Cys and GAG for residu  
A:Note: this translation is not annotated in GenBank entries HSWRSX1A, HSWRSX1B, HSWR  
R:Frolova, L.Y.; Sudomoina, M.A.; Grigorjeva, A.Y.; Zinovjeva, O.L.; Kisseliev, L.L.  
Gene 109, 291-296, 1991  
A:Title: Cloning and nucleotide sequence of the structural gene encoding for human tr  
A:Reference number: JH0533; MUID:92112058  
A:Accession: JH0533  
A:Molecule type: mRNA



;;  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: SmithKline Beecham Corporation  
;; STREET: 709 Swedeland Road  
;; CITY: King of Prussia  
;; STATE: PA  
;; COUNTRY: USA  
;; ZIP: 19406-0939  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Diskette  
;; COMPUTER: IBM Compatible  
;; OPERATING SYSTEM: DOS  
;; SOFTWARE: FASTSEQ for Windows Version 2.0  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/928,100  
;; FILING DATE: 12-SEP-1997  
;; CLASSIFICATION: 424  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 9619072.3  
;; FILING DATE: 12-SEP-1996  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Gimmil, Edward R  
;; REGISTRATION NUMBER: 38,891  
;; REFERENCE/DOCKET NUMBER: P31624-1  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 610-270-4478  
;; TELEFAX: 610-270-5090  
;; TELEX:  
;; INFORMATION FOR SEQ ID NO: 4:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 197 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
;; US-08-928-100-4  
;  
Query Match 4.2%; Score 108; DB 3; Length 197;  
Best Local Similarity 25.3%; Pred. No. 0.0016;  
Matches 42; Conservative 27; Mismatches 73; Indels 24; Gaps 6;  
;  
QY 307 IPCAIDDDPYFRMTRD-----VAPRIGYKPKALLSTFPALOGNOTKMSA 352  
DB 5 VPGVTDQKPMIEQREIVRSFNNAVNCVDVLEPEGIYPENE--RAGRLPGIDG-NAKMSK 61  
;  
QY 353 SDPNSSIFLTDPAKQITKYKNAHAFSGGRDIEHRQFGNCVDVSPMYLTFF--LEDD 410  
DB 62 S-LNNGIYLDADDTLKKKWSMTDPDHIRVEDPGKIEGN---WPHYILDVGFREDA 116  
;  
QY 411 DKLEQIRKDYTSGAMLTGELKALIEVLOPLIAEHQARRKEVTDEI 456  
DB 117 QEIDMKERYGRGIGDVKTKRYLLEIRLELGFIRERRIEFKADM 162  
;  
RESULT 13  
US-09-183-134-2  
; Sequence 2, Application US/09183134  
; Patent No. 6165759  
; GENERAL INFORMATION:  
; APPLICANT: Lawlor, Elizabeth  
; TITLE OF INVENTION: NOVEL TRYPTOPHANYL TRNA SYNTHETASE  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SmithKline Beecham Corporation  
; STREET: 709 Swedeland Road  
; CITY: King of Prussia  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19406-0939  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FASTSEQ for Windows Version 2.0

;;  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/09/183,134  
;; FILING DATE:  
;; CLASSIFICATION:  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 08/923,867  
;; FILING DATE: 04-SEP-1997  
;; APPLICATION NUMBER: 9619072.3  
;; FILING DATE: 12-SEP-1996  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Gimmil, Edward R  
;; REGISTRATION NUMBER: 38,891  
;; REFERENCE/DOCKET NUMBER: P31624  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 610-270-4478  
;; TELEFAX: 610-270-5090  
;; TELEX:  
;; INFORMATION FOR SEQ ID NO: 2:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 197 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
;; US-09-183-134-2  
;  
Query Match 4.2%; Score 108; DB 4; Length 197;  
Best Local Similarity 25.3%; Pred. No. 0.0016;  
Matches 42; Conservative 27; Mismatches 73; Indels 24; Gaps 6;  
;  
QY 307 IPCAIDDDPYFRMTRD-----VAPRIGYKPKALLSTFPALOGNOTKMSA 352  
DB 5 VPGVTDQKPMIEQREIVRSFNNAVNCVDVLEPEGIYPENE--RAGRLPGIDG-NAKMSK 61  
;  
QY 353 SDPNSSIFLTDPAKQITKYKNAHAFSGGRDIEHRQFGNCVDVSPMYLTFF--LEDD 410  
DB 62 S-LNNGIYLDADDTLKKKWSMTDPDHIRVEDPGKIEGN---WPHYILDVGFREDA 116  
;  
QY 411 DKLEQIRKDYTSGAMLTGELKALIEVLOPLIAEHQARRKEVTDEI 456  
DB 117 QEIDMKERYGRGIGDVKTKRYLLEIRLELGFIRERRIEFKADM 162  
;  
RESULT 14  
US-09-492-581-4  
; Sequence 4, Application US/09492581  
; Patent No. 6346409  
; GENERAL INFORMATION:  
; APPLICANT: Gentry, Danile  
; APPLICANT: Greenwood, Claire  
; APPLICANT: Lawlor, Elizabeth  
; TITLE OF INVENTION: No. 6346409el trps  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SmithKline Beecham Corporation  
; STREET: 709 Swedeland Road  
; CITY: King of Prussia  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19406-0939  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FASTSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/492,581  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/928,100  
; FILING DATE: 12-SEP-1997  
; APPLICATION NUMBER: 9619072.3

```

Db 153 S--RLSSGISFTEFTYOILQSIDFY--TLAKKHNIOLQIGADOWGNTAGLILIRKKE 207
QY 266 -----JEGFP-----DSDCIGKISFPALQAPSPNSPP--QIRPDRDIOCLIPCAI 311
Db 208 GPEAKVFGLTIPMLKADGKFGTAGATWLDPKKTSPEFYQFWINQDD----- 258
QY 312 DDDPYFRMPDVAIRIGYPKAPALHSTFPALQAGQTKMSASDPNSSIFLTDTAKOIKTK 371
Db 259 -----RDV---IKYLK-----FFTFDLKEE-----IDALAEKYEKE 286
QY 372 VKNHANSGGDTTIEHRQFGGNCDDVVSFMYLTFLEDDDKLEQIKRDYTGAMLTGELK 431
Db 287 PKRR--EAQRRLAE-----VTFVHDDALEBAOK--ISEALFSGNIK 326
QY 432 KALIEVQ-----PLIAHQARRKEVTDIVKEEPTPK 465
Db 327 DLITEIEGLENHVPY-ELTKDAKNIVDWLVDTIEPSK 365

```

```

RESULT 10
US-09-352-990-28
; Sequence 28, Application US/09352990
; Patent No. 6255090
; GENERAL INFORMATION:
; APPLICANT: Famodu, Layo O.
; APPLICANT: Orozco, Buddy
; APPLICANT: Rafalski, Antoni
; TITLE OF INVENTION: Plant Aminoacyl-tRNA Synthetase
; FILE REFERENCE: BR-1191
; CURRENT APPLICATION NUMBER: US/09/352, 990
; CURRENT FILING DATE: 1999-07-14
; EARLIER APPLICATION NUMBER: 60/092, 866
; EARLIER FILING DATE: July 15, 1998
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 28
; LENGTH: 377
; TYPE: PRT
; ORGANISM: Synecchocystis sp.
US-09-352-990-28

```

Query Match 4.3%; Score 109; DB 4; Length 377;  
 Best Local Similarity 21.0%; Pred. No. 0.0037;  
 Matches 82; Conservative 49; Mismatches 116; Indels 144; Gaps 19;

```

QY 138 FSHRDMNQVLDAYENKRFYLTGRGSSSEAMVGHILPFIPTKQLQDVENVLYIQMTD 197
Db 33 FAHSTTAM-----DKPRLT-SGVOPTGN-LHLGNLTGAI-RSWWQO-----QOHY 75
QY 198 DEKYLWKDL-----TLDOAYGDAVENAKDILACGFDINKTFTF----- 235
Db 76 DNEFCVVDLAIYVPHNPQLAQ---DTLITIALYIACGIDLOXSTIFVQSHVAHSELA 132
QY 236 -----SDLDYMGNSGSEFYKVNVIQKHVTFNQVKGITGFTSDSCIGKISFPALQAPSF 289
Db 133 WLLNCVYPLMWLERMIOFKKAVKQGENVS-----VCLLDYPLVMAA--- 174
QY 290 SNSEFQIFRRRTIOCL-----IPCALDQDPYFRMTRDV-----APRIGYKPA 333
Db 175 -----DILLYDADKVPREGQKQHLTTRDVIIRINDKREGEDAPVLIKLEPL 222
QY 334 L-LHSTFPALQAGQTKMSADPN--SSIFLTDTAKOIKTKVNKHAFFSGGRDTEEHROF 390
Db 223 IRKEGAVMSLADGTTKMSKMSDESELSRINLLDPEMIKKKKVKK----- 266
QY 391 GNCDDVVSFMYLTFLEDDDKLEQIKRDYTGAMLTGELKKAL----- 434
Db 267 ---CKTDPQ--RGTMF---DDPERECHNLLTYLLISNQTKEAVDQCAEMGQCFKPL 319
QY 435 -----IEVLOPLIAEHQARRKEVTDIVK 458
Db 320 TETAIAALEPIQAKYAEITLADRGEL-DRITQ 349

```

```

RESULT 11
US-08-923-867-2
; Sequence 2, Application US/08923867
; Patent No. 5851809
; GENERAL INFORMATION:
; APPLICANT: Lawlor, Elizabeth
; TITLE OF INVENTION: NOVEL TRYPTOPHANYL tRNA
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/923, 867
; FILING DATE: 04-SEP-1997
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 9619072.3
; FILING DATE: 12-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimmil, Edward R.
; REGISTRATION NUMBER: 38, 891
; REFERENCE/DOCKET NUMBER: P31624
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 197 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-923-867-2

```

Query Match 4.2%; Score 108; DB 2; Length 197;  
 Best Local Similarity 25.3%; Pred. No. 0.0016;  
 Matches 42; Conservative 27; Mismatches 73; Indels 24; Gaps 6;

```

QY 307 IPCALDQDPYFRMTRD-----VAPRIGYKPPALHSTFPALQAGQTKMSA 352
Db 5 VPGTQDKPMIEDQTRVRSFNNAYNQDVLYVEPEGITPENE--RAGRLPQLDG-NKMKK 61
QY 353 SDPNSSIFLTDTAKOIKTKVNKHAFFSGGRDTEEHROFGGNCDDVVSFMYLTF--LEDD 410
Db 62 S-LNNGIYIADDDTLRKVKVSMYTPDPIRYVDPKIEGN---WVFHYLDVFGRPEDA 116
QY 411 DKLEQIKRDYTGAMLTGELKKALIEVLOPLIAEHQARRKEVTDIVK 456
Db 117 QETADMKERYQRGGLGDVTKRYLLELTELRLGPIRRRIEFARKM 162

```







US-09-492-581-2

Query Match	7.3%;	Score 185.5;	DB 4;	Length 341;
Best Local Similarity	24.0%;	Pred. No. 1.4e-11;		
Matches	81;	Conservative	60;	Mismatches 128;
			Indels	69;
			Gaps	15;

```

OY 154 KPEFLYUWGRSSAMVHGHILPIEFKMLDYDENVNVLATOMTDKXY-LMKDITDPOAY 212
Dp 3 KPIIL-TGDRTGK-LHIGHVYSLKRR-----VLOEDKXMDYFVLEADQAL 49
OY 213 GDAVEN-----AKDIACGDFINKTF--SDLDGMGSGCFKNV---KIQ 256
Dp 50 TDHAKDPOTIVESIGNVALDYLAGDLPNNSTJFIQDPELWELSYMYNVLSTARLER 109
OY 257 HATFNQWKGIFGFTSDICIKISPAIOQA--PFSNSFPQIFRDRDIOCLIPCAIDOD 314
Dp 110 NPTVTEISQKGFESIPFTGFLVYRIQADDTAFKANY-----VPVGTQK 156
OY 315 PYFKMTD-----VAPRIGYKPRALLHSTFFPALQOAGOTKMSADPNSSIF 360
Dp 157 PMIEOTREIYNSFNNAVNCVULVEPEGIYEPNE--RAGRLEPGIDG-NAKMSKS-LNNGIY 212
OY 361 LTRPAKOIKTVKNHAFGSGHDTIEEHRQSGMCDVUVSEMYULFF--LEDDDKLEIRK 418
Dp 213 LADADLTLRKVVMSMYDPRHIVREDPGKIEGN---MYVNIYDUFEREPDQAEIADMKE 268
OY 419 DYTSGAMITGELKALLLEVLOPLAEHQARKKEVTDEI 456
Dp 269 RYQRGGLADVYTKRYLLEITERELGPIRRERIEPAKOM 306

```

RESULT 4  
US-08-743-130A-39  
; Sequence 39, Application US/08743130A

```

: GENERAL INFORMATION:
: APPLICANT: Sassanfar, Mandana
: APPLICANT: Gallant, Paul L.
: APPLICANT: Shen, Xiaoyu
: APPLICANT: Tao, Nianjun
: APPLICANT: Tao, Jianshi
: APPLICANT: Houman, Fariba
: TITLE OF INVENTION: CANDIDA TYROSYL-1RNA SYNTHETASE
: TITLE OF INVENTION: PROTEINS, NUCLEIC ACIDS AND STRAINS COMPRISING SAME
: NUMBER OF SEQUENCES: 41

```

```

COMPUTER READABLE FORM :
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/743,130A

```

ATTORNEY/AGENT INFORMATION:  
NAME: Brook Esq., David E.  
REGISTRATION NUMBER: 22,592  
REFERENCE/DOCKET NUMBER: CPT195-12  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 861-6240  
TELEFAX: (617) 861-9540  
INFORMATION FOR SEQ. ID NO.: 39:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 409 amino acids  
TYPE: amino acid  
STRANDEDNESS:

```

;          TOPOLOGY: linear
US-08-743-130A-39

```

Query Match	6.7%;	Score 169.5;	DB 2;	Length 409;
Best Local Similarity	20.6%;	Pred. No. 1.1e-09;		
Matches 79; Conservative	61;	Mismatches 136;	Indels 107;	Gaps 16;

```

QY 146 VLDAYENK-KPEVLYTORGPSESMHNGHILPFI-----FTKWLQDV-----F 187
Db 27 IKDVLKENRPVKYITWGSTARPGKR-HCGYFVPMKILAHFLKAGCEYVLLADLHAFLDNM 85
QY 188 NVPLVIOMTDDEKYLKMDLTLDDQYSGDAVENAKDIIACGEDINKFTFESLDYWGMSGGF 247
Db 86 KAPLEVVKYRAKAYEEFVVKALIKSINVPIERLKFVWGSSYQKGGDY---MDLEFKLSINV 142
QY 248 YKNVVK-----IQCHVTFPNQYKGFIFGTTSDSCICKISFPALQAAPEFSNSFPQIFRRTD 302
Db 143 SQNDAKRAGADVNVQVANPLISGLI-----YPLMQA-----IDEEHLG 180
QY 303 IQCLIPALIDDDPEFRMTDVAPEIIGYKPRALLHSTFEPFALOGAQOTKMSADPPSSIFLT 362
Db 181 VDAOFG-GVDDRKFTFLAEENLPISITGKRAHLMNPMVPGI-GGGKMSASDPNSKIDIT 238
QY 363 DTAQIKITKVKNAHAFSGG--RDT-----IEEHROFGCN 393
Db 239 EEPKVVKKVNSAIVCAHGGELKNDNLIAFIEVIOPIALFKTGVGCAFKDIDREKKGCG- 297
QY 394 CDVDSVFMYLTFLEDDDKLEQIRKDYTSGAMLTGELKKALI---EVLOPIAEHQARR 449
Db 298 ---PLSY-----DSIQRLKADEVGDKLAPPLDKSGVADKINELLAPIRAEFESS- 343
QY 450 KEVYDVLKFEFMPRLKLSPDFOK 472
Db 344 -----EEFQVAKNGKGYVEK 358

```

RESULT 5  
US-08-743-130A-2  
; Sequence 2, Application US/08743130A  
; Patent No. 5871987

1 APPLICANT: Sassanfar, Mandana  
 2  
 3 APPLICANT: Gallant, Paul L.  
 4  
 5 APPLICANT: Shen, Xiaoyu  
 6  
 7 APPLICANT: Tao, Mianjun  
 8  
 9 APPLICANT: Tao, Jianshi  
 10  
 11 APPLICANT: Houman, Fariba  
 12  
 13 TITLE OF INVENTION: CANDIDA TYROSYL-TRNA SYNTHETASE  
 14  
 15 TITLE OF INVENTION: PROTEIN, NUCLEIC ACIDS AND STRAINS COMPRISING SAME  
 16  
 17 NUMBER OF SEQUENCES: 41  
 18  
 19 CORRESPONDENCE ADDRESS:

```

? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? OPERATING SYSTEM: IBM PC compatible
? SOFTWARE: Patent In Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/743,130A

```

CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Brook Esq., David E.  
REGISTRATION NUMBER: 22,592  
REFERENCE/DOCKET NUMBER: CP195-12  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 861-6240  
TELEFAX: (617) 861-9540

OY	134	RGIFESHEDHNOQVIDAENKPKFELLYTGRGSSSAMHGHILPIETFKMLQDVENPLVI	193
OY	134	RGIFESHEDHNOQVIDAENKPKFELLYTGRGSSSAMHGHILPIETFKMLQDVENPLVI	193
Db	73	RGVFESQORDDIRIIDLYEHGHEFFLYTGRGSSSMHGHMVPITFKMLQAEVDPPLVI	132
OY	194	QMTDEKYLAK-DLTLDOAGDAVENAKDIACGFDINKTPIESDLYMGMSSEGYKNV	252
Db	133	ELTDEKELFHHOITLIDVKKCFKAENNAKDIAGFENBENPIESDLOYMC--CAFEYENV	190
OY	253	KIOKHVFNQYKGFIEGFUDSCIOKISFPALQAPSENSPPOFRDRTIOCLIPACID	312
Db	191	RTRQIITSTAKANVGFGLDSOICKINHASIQIINFATPSSPDVLGLPRTICLIPACID	250
OY	313	QDPYFRMTRDVAPRIQYKPPALHSTFPPALQOAKTMSADPNSSIFLTDIAKIKTV	372
Db	251	QDPYFRMTRDVADRIKRFKPPALIHAKFFPALQOASTKMSADTTTSIFMSDITAKIOKIKI	310
OY	373	NKHAFFSGGRDYTEEHROPGGNDVDVSEMYLTFLEDDDKLEQIRKYDTSGAMLTGLKX	432
Db	311	NKYAFSGGRATAEHEHRELIGCNPEVDVAFQYLSEFSYDDEKALQOLEGYRRGELLSGMKX	370
OY	433	ALIEVLOPLIAHQAARKEVTDVETAYKEMPRKLSF	468
Db	371	ECITVLOEVSAYQERRSKVDQVVEKFEKHKILVE	406

RESULT 2  
US-08-928-100-2  
; Sequence 2, Application US/08928100

```

1  GENERAL INFORMATION:
2  APPLICANT: Gentry, Danile
3  APPLICANT: Greenwood, Claire
4  APPLICANT: Lawlor, Elizabeth
5  TITLE OF INVENTION: No. 6046174e1 tips
6  NUMBER OF SEQUENCES: 6
7  CORRESPONDENCE ADDRESS:
8  ADDRESSEE: Smlthklne Beecham Corporation
9  STREET: 709 Swedeland Road
10 CITY: King of Prussia
11 STATE: PA
12 COUNTRY: USA
13 ZIP: 19406-0939
14 COMPUTER READABLE FORM:
15 MEDIUM TYPE: Diskette
16 COMPUTER: IBM Compatible
17 OPERATING SYSTEM: DOS
18 SOFTWARE: FastSeq for Windows Version 2.0
19 CURRENT APPLICATION DATA:
20 APPLICATION NUMBER: US/08/928,100
21 FILING DATE: 12-SEP-1997
22 CLASSIFICATION: 424
23 PRIOR APPLICATION DATA:
24 APPLICATION NUMBER: 9619072.3
25 FILING DATE: 12-SEP-1996
26 ATTORNEY/AGENT INFORMATION:
27 NAME: Gimm, Edward R
28 REGISTRATION NUMBER: 38,891
29 REFERENCE/DOCKET NUMBER: P31624-1
30 TELECOMMUNICATION INFORMATION:
31 TELEPHONE: 610-270-4478
32 TELEFAX: 610-270-5090
33 TELEX:
34 INFORMATION FOR SEQ ID NO: 2:
35 SEQUENCE CHARACTERISTICS:
36 LENGTH: 341 amino acids
37 TYPE: amino acid
38 STRANDEDNESS: single
39 TOPOLOGY: linear
40 MOLECULE TYPE: protein
41 OS-08-928-100-2

```

Query Match	7.3%;	Score 185.5;	DB 3;	Length 341;
Best Local Similarity	24.0%;	Pred. No. 1.4e-11;		

[illegible]

RESULT 3  
US-09-492-581-2

```

1 GENERAL INFORMATION:
2 APPLICANT: Gentry, Danlle
3 APPLICANT: Greenwood, Claire
4 APPLICANT: Lawlor, Elizabeth
5 TITLE OF INVENTION: No. 6346409e1 trps
6 NUMBER OF SEQUENCES: 6
7 CORRESPONDENCE ADDRESS:
8 ADDRESSEE: SmithKline Beecham Corporation
9 STREET: 709 Swedeland Road
10 CITY: King of Prussia
11 STATE: PA
12 COUNTRY: USA
13 ZIP: 19406-0939
14 COMPUTER READABLE FORM:
15 MEDIUM TYPE: Diskette
16 COMPUTER: IBM Compatible
17 OPERATING SYSTEM: DOS
18 SOFTWARE: FASTSEQ for Windows Version 2.0
19 CURRENT APPLICATION DATA:
20 APPLICATION NUMBER: US/09/492,581
21 FILING DATE:
22 CLASSIFICATION:
23 PRIORITY APPLICATION DATA:
24 APPLICATION NUMBER: US/08/928,100
25 FILING DATE: 12-SEP-1997
26 APPLICATION NUMBER: 9619072.3
27 FILING DATE: 12-SEP-1996
28 ATTORNEY/AGENT INFORMATION:
29 NAME: Gimm1, Edward R
30 REGISTRATION NUMBER: 38,891
31 REFERENCE/DOCKET NUMBER: P31624-1
32 TELECOMMUNICATION INFORMATION:
33 TELEPHONE: 610-270-4478
34 TELEFAX: 610-270-5090
35 TELEEX:
36 INFORMATION FOR SEQ ID NO: 2:
37 SEQUENCE CHARACTERISTICS:
38 LENGTH: 341 amino acids
39 TYPE: amino acid
40 STRANDEDNESS: single
41 TOPOLOGY: linear
42 MOLECULE TYPE: protein

```

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 24, 2002, 12:49:45 ; Search time 11.8319 Seconds  
(Without alignments)  
999.159 Million cell updates/sec

Title: US-09-813-718-10  
Perfect score: 2547  
Sequence: 1 MPNSEPASLELNFNSTATQC.....KLSEDFOKLALEHHHHHH 484

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :  
1: Issued\_Patents\_AA:\*  
2: /cgn2\_6/ptodata/1/1aa/5A-COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/5B-COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6A-COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/6B-COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/6CTUS-COMB.pep:\*  
7: /cgn2\_6/ptodata/1/1aa/6CTUS-COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1218.5	47.8	424	US-08-876-885-26	Sequence 26, Appl
2	185.5	7.3	341	US-08-928-100-2	Sequence 2, Appl
3	185.5	7.3	341	US-09-492-581-2	Sequence 2, Appl
4	169.5	6.7	409	US-08-743-130A-39	Sequence 39, Appl
5	167.5	6.6	409	US-08-743-130A-2	Sequence 2, Appl
6	153	6.0	1440	US-09-357-251-37	Sequence 37, Appl
7	137	5.4	388	US-08-705-868-4	Sequence 4, Appl
8	137	5.4	388	US-09-123-615-4	Sequence 4, Appl
9	113.5	4.5	418	US-08-855-910-11	Sequence 11, Appl
10	109	4.3	377	US-09-352-990-28	Sequence 28, Appl
11	108	4.2	197	US-08-923-867-2	Sequence 2, Appl
12	108	4.2	197	US-08-928-100-4	Sequence 4, Appl
13	108	4.2	197	US-09-183-134-2	Sequence 2, Appl
14	108	4.2	197	US-09-492-581-4	Sequence 4, Appl
15	104.5	4.1	370	US-08-415-593-45	Sequence 45, Appl
16	95	3.7	323	US-08-816-977-21	Sequence 21, Appl
17	95	3.7	418	US-08-844-054-2	Sequence 2, Appl
18	95	3.7	418	US-09-347-333-2	Sequence 2, Appl
19	94.5	3.7	877	US-08-907-166-8	Sequence 8, Appl
20	94	3.7	855	US-08-890-865A-10	Sequence 10, Appl
21	90.5	3.6	898	US-08-465-995A-4	Sequence 4, Appl
22	90.5	3.6	898	US-08-465-994C-4	Sequence 4, Appl
23	90.5	3.6	898	US-08-966-145-4	Sequence 4, Appl
24	90.5	3.6	920	US-08-101-593-4	Sequence 4, Appl
25	89.5	3.5	344	US-09-393-554-2	Sequence 2, Appl
26	89.5	3.5	898	US-08-465-995A-2	Sequence 2, Appl
27	89.5	3.5	898	US-08-465-994C-2	Sequence 2, Appl

28	89.5	3.5	898	2	US-08-966-145-2	Sequence 2, Appl
29	89.5	3.5	920	1	US-08-101-593-2	Sequence 2, Appl
30	89	3.5	719	1	US-08-082-849B-31	Sequence 31, Appl
31	89	3.5	719	5	PCT-US94-0162A-31	Sequence 31, Appl
32	89	3.5	2206	1	US-07-852-260-2	Sequence 2, Appl
33	89	3.5	2206	1	US-08-461-503-2	Sequence 2, Appl
34	89	3.5	2206	4	US-08-465-250-2	Sequence 2, Appl
35	88	3.5	606	2	US-08-883-534-3	Sequence 3, Appl
36	88	3.5	606	3	US-09-204-764-3	Sequence 3, Appl
37	87	3.4	428	3	US-08-331-625A-43	Sequence 43, Appl
38	87	3.4	1452	3	US-08-331-625A-2	Sequence 2, Appl
39	87	3.4	1452	5	PCT-US93-04384-18	Sequence 18, Appl
40	87	3.4	1452	5	PCT-US93-04692-2	Sequence 2, Appl
41	86.5	3.4	495	4	US-09-217-490-2	Sequence 2, Appl
42	86	3.4	1443	1	US-08-308-872B-2	Sequence 2, Appl
43	85	3.3	410	4	US-09-352-990-16	Sequence 16, Appl
44	83.5	3.3	510	1	US-08-249-112-3	Sequence 3, Appl
45	83.5	3.3	510	4	US-09-522-217-89	Sequence 89, Appl

## ALIGNMENTS

RESULT 1  
US-08-876-885-26  
Sequence 26, Application US/08876885  
Patent No. 6174713  
GENERAL INFORMATION:  
APPLICANT: Shen, Xiaoyu  
APPLICANT: Houman, Fatiba  
TITLE OF INVENTION: CANDIDA CYTOPLASMIC TRYPTOPANYL-tRNA  
TITLE OF INVENTION: SANTHEPASE PROTEINS, NUCLEIC ACIDS AND STRAINS COMPRISING  
NUMBER OF INVENTION: SAME  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
STREET: Two Millitia Drive  
CITY: Lexington  
STATE: MA  
COUNTRY: USA  
ZIP: 02173  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/876,885  
FILING DATE: 16-JUN-1997  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Brook, David E.  
REGISTRATION NUMBER: 22,592  
REFERENCE/DOCKET NUMBER: CPI97-02  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (781) 861-6240  
TELEFAX: (781) 861-9540  
INFORMATION FOR SEQ ID NO: 26:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 424 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: Linear  
US-08-876-885-26

Query Match 47.8%; Score 1218.5; DB 4; Length 424;  
Best Local Similarity 57.1%; Pred. No. 6.3e-124;  
Matches 226; Conservative 71; Mismatches 92; Indels 7; Gaps 3;

QY 78 TEAEDEPDEPTVQ---TSSAKGIDYKLVGRSSKIDKELNRIERATGCRPHHFLR 133  
DB 13 TESEDKTPWEVGCAVVDGKSMGIDYDKLISOPCTKHITETLEFRKQVNGEHPFLK 72

CC abyss1 (see AAF86431 and AAH41223-7) and P. abyss1 proteins. P. abyss1 is  
 CC a hyperthermophilic archaeon, which is isolated from deep-sea  
 CC hydrothermal vents. The present sequence is one such P. abyss1 protein.  
 CC The proteins of the present invention have various potential industrial  
 CC uses, since the proteins are stable at very high temperatures, some up to  
 CC 110 degrees centigrade.  
 CC Note: This patent is in the same patent family as WO200065062, which  
 CC contains additional sequences as shown in AAB99132-AAB99143,  
 CC AAH75903-AAH75920 and AAG66436.  
 CC  
 XX  
 SQ Sequence 385 AA:  
 Query Match 31.5%; Score 803; DB 22; Length 385;  
 Best Local Similarity 45.3%; Pred. No. 2,2e-71;  
 Matches 173; Conservative 63; Mismatches 120; Indels 26; Gaps 9;  
 QY 82 EDFVDPWVQTSANGIDYDKLIVRGSSKIDKELINRIERATGQRPHHFLRGIFESH 140  
 DB 3 EDFKVPWEVEGV---VDYKNKLIENHGTSPLTEELLEKTAELTKSELPLFFPRKFFESH 58  
 QY 141 RDMNOYLDAYENKKPFYLYTGRGPSSAMHVGHLIPFTFKWLODVFNPLVQMTDDEK 200  
 DB 59 RQDKVLYQDYEGRGFLYTGRRPGSG-PMHIGHTLPEFATKWLQEKFGVNLVQITDDEK 117  
 QY 201 YLWMD-LTLDQAYGDAVENAKDIACGFDINKTFIFSDLYMGSSGFYKNVVKIQKHVT 259  
 DB 118 FLTKENITFEEDTKHMAVENLDIIAVGFDDPKTFIPQNSEF---ITLYEMALPIAKKIN 173  
 QY 260 FNOYKGIPTGTDSDCIKISFPAIQAAFSNSFPQIFRDRDIOCLIPCAIDDDPYFM 319  
 DB 174 FSKAKVFGFTGEOISKIGMIFFPALQIAPTF-----FEKR---RCLIPAIQDDPYWRL 223  
 QY 320 TROVAPRIGYKPKALLHSTFPAQLQGMQTKMSADPNSSIFLIDTAKQITKYNKHAFSG 379  
 DB 224 QRPFAESLGYKTAALHSKFPVSLTSLSGKMSASKPETALVLTDSPEVKKWKWKEALTG 283  
 QY 380 GRDTIEHRQFGNCDVDSFMYLTFLEDDDKLEQIRKDY---TSGAMLTGELKALIE 436  
 DB 284 GRFLKEQRRKGGEPEKCVFKMLEIFFEEDDK--KLKERYACKNGNELTCGCKRYLIS 341  
 QY 437 VLQPLIAHQARKKQVTDITYK 458  
 DB 342 KIOEFLKEHQRRKKAQIEK 363  
 RESULT 15  
 AAB58219  
 ID AAB58219 standard; Protein; 142 AA.  
 XX  
 AC AAB58219;  
 XX  
 DT 14-MAR-2001 (first entry)  
 XX  
 DE Lung cancer associated polypeptide sequence SEQ ID 557.  
 XX  
 KM Human; Lung cancer associated protein; neuroprotective; cytosstatic;  
 KM cardioactive; immunomodulatory; muscular active; vulnerrary;  
 KM gastrointestinal; nephroproptic; antinfecitive; gynecological;  
 KM antibacterial; diagnosis; neural disorder; immune disorder; reproductive;  
 KM proliferative disorder; wound healing; infectious disease.  
 XX  
 OS Homo saplens.  
 XX  
 PN WO200055180-A2.  
 PD  
 XX 21-SEP-2000.  
 PF  
 XX 08-MAR-2000; 2000WO-US05918.  
 PR 12-MAR-1999; 99US-0124270.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 PA (ROSE/) ROSEN C A.

XX  
 PI Ruben SM;  
 XX  
 DR WPI; 2000-587514/55.  
 DR N-PSDB; AAF18095.  
 XX  
 PT Lung cancer associated gene sequences, referred to as lung cancer  
 PT antigens, useful for treatment, prevention, and diagnosis of disorders  
 PT such as lung cancer -  
 XX  
 PS Claim 11; Page 1051; 1425pp; English.  
 XX  
 CC Polynucleotide sequences AAF17982 - AAF18424 encode human lung cancer  
 CC associated proteins represented in AAB58106 - AAB58548. Lung cancer  
 CC associated proteins and polynucleotide sequences, their agonists, and  
 CC antagonists may have neuroprotective; cytosstatic; cardioactive;  
 CC immunomodulatory; muscular active general; vulnerrary; gastrointestinal  
 CC general; nephroproptic; antinfecitive; gynecological; or antibacterial  
 CC activity. The invention also includes antibodies specific for the  
 CC protein or polynucleotide sequences. The lung cancer associated  
 CC polynucleotide sequences may be used for detection of lung cancer,  
 CC chromosome identification, as chromosome markers, and for numerous other  
 CC diagnostic or research purposes. The proteins may be used to treat  
 CC disorders such as neural, immune, muscular, reproductive,  
 CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative  
 CC disorders. The proteins may also be used in the treatment of wounds and  
 CC infectious diseases. Polynucleotide sequences AAF18425 - AAF18433 and  
 CC peptide AAB58349 are used in the course of the invention for the  
 CC identification and characterisation of the polynucleotide and protein  
 CC sequences.  
 XX  
 SQ Sequence 142 AA:  
 Query Match 22.8%; Score 579.5; DB 21; Length 142;  
 Best Local Similarity 85.5%; Pred. No. 1.1e-49;  
 Matches 118; Conservative 5; Mismatches 12; Indels 3; Gaps 2;  
 QY 1 MPNSEPASLLEFNSTATOGELVRSIKAGNASDEIDSAVKMLVSLKMSYKAAAGDYKA 60  
 DB 5 MPNSEPASLLEFNSTATOGELVRSIKAGNASDEIDSAVKMLVSLKMSYKAAAGDYKA 64  
 QY 61 DCPGPNAPTSMHGPDAEEDFVDPWVQTSANGIDYDKLIVRGSSKIDKELINRI 120  
 DB 65 DCPGPNAPTSMHGPDAEEDFVDPWVQTSANGIDYDKLIVRGSSXNNQELLXDX 124  
 QY 121 ERATGQRPHHFLRGITFF 138  
 DB 125 E-STAKXTHS--GQGXFF 139

Search completed: October 24, 2002, 12:51:13  
 Job time : 33.3193 secs



PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match	38.0%	Score 968	DB 21	Length 292
Best Local Similarity	60.9%	Pred. No. 4,4e-88		
Matches 179; Conservative	51	Mismatches 62	Indels 2	Gaps 1

QY	178	IFETWIDOVNVPVLTOMTDEKXIMKMDTLDDQXGSAVBNADIIACGSDIJKTFEED	237
Db	1	MEFIILOAEFVPLVLTOTDEKSIKMLNLSIESQRLAREAKOIIACGFDVITKTFEED	60
QY	238	LDYMGMSGFKNVVKLOKHNTFNOVQVGIGETTSDICIGTISPALIOAAFSFNSPQIF	297
Db	61	FDVYG - GAFYKNVYKGVCKTILKAMGIFGFSGEDIIAKLSPFPVQVAFSPSSPHLF	118
QY	298	RDRDIOCLIFCALDOPPYFMTDVAVPRIGYKPKPALHSTFFPALOGAOTKMSASPNS	357
Db	119	PGKNLRCILPICALDOPPYFMTDVAVPRIGYSKPALISTFFPALOGENGKMSASPNS	178
QY	358	SIFLTFPAKOJKTUVKNHAFSGSGRDTIEEHKFGSGNCDVAFSEWYLFLEDDDKLBOIR	417
Db	179	AIYTTDSAKDIOKNNINRYAFSGGDSITEKHRELGCANLEVOIIPVKYLSFLEDDSELEHIK	238
QY	418	KDYTSGAMLTGELKALIEVLOPLIAEHQARRKEVTEIYKEFMTPKLSDFQ	471
Db	239	KEYEGRMLTGVKKRLTEVTEIYEHKRRRAAVTENDAFVAPVPSKE	292

RESULT 13	
AAB58517	
ID	AAB58517 standard; Protein; 173 AA

XX	14-MAR-2001	(first entry)
XX		
XX		
XX		
DE	Lung cancer associated polypeptide sequence SEQ ID 855	

KM Human; lung cancer associated protein; neuroprotective; cytostatic;  
KM cardioactive; immunomodulatory; muscular active; vulnerary;  
KM gastrointestinal; nephrotropic; antiinfective; gynecological;  
KM antibacterial; diagnosis; neural disorder; immune disorder; reproductive  
KM proliferative disorder; wound healing; infectious disease.

OS	Homo sapiens.
XX	
PN	W020005180-A2.
XX	
PD	21-SEP-2000.
XX	
PE	08-MAR-2000; 2000OWO-US05918.
XX	
PR	12-MAR-1999; 9905-0124270.
XX	
PA	(HUMA-) HUMAN GENOME SCI INC.
PA	(ROSE/ ROSEN C A.

AA	
PI	Ruben SM;
XX	
DR	WPI; 2000-587514/55.
DR	N-PSDB; AAF18393.

PT Lung cancer associated gene sequences, referred to as lung cancer  
PT antigens, useful for treatment, prevention, and diagnosis of disorders  
PT such as lung cancer .

Claim 11; Page 1392-1393; 1425pp; English.

Polynucleotide sequences AAF17982 - AAF18424 encode human Lung cancer

associated proteins represented in AAB58106 - AAB58348. Lung cancer associated proteins and polynucleotide sequences, their agonists, and antagonists may have neuroprotective; cytostatic; cardiocactive; immunomodulatory; muscular active general; vulnary; gastrointestinal general; nephrotoxic; antineoplastic; gynecological; or antibacterial activity. The invention also includes antibodies specific for the protein or polynucleotide sequences. The lung cancer associated polynucleotide sequences may be used for detection of lung cancer, chromosome identification, as chromosome markers, and for numerous other diagnostic or research purposes. The proteins may be used to treat disorders such as neural, immune, muscular, reproductive, gastrointestinal, pulmonary, cardiovascular, renal, and proliferative disorders. The proteins may also be used in the treatment of wounds and infectious diseases. Polynucleotide sequences AAF18425 - AAF18433 and peptide AAB58349 are used in the course of the invention for the identification and characterisation of the polynucleotide and protein sequences.

Seq	Sequence	173 AA;
Query Match	32.6%;	Score 831; DB 21; Length 173;
Best Local Similarity	98.7%;	Pred. No. 1e-74;
Matches 156;	Conservative 0;	Mismatches 2; Indels 0; Gaps 0

[illegible]

**Qy**    237 DLDYMGSSSGFYKNVYKIQKHVTFNQYKGIFGFTDSDCGIKISFPAIQAPSFNSNFPQI 296  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
**Db**    61 DLDYMGSSSGFYKNVYKIQKHVTFNQYKGIFGFTDSDCGIKISFPAIQAPSFNSNFPQI 120

QY	297	FRDRFDIOCLIPCAIDDPYFRMTRDVAPRIGYKPAL	334
Db	121	FRDRFDIOCLIPCAIDDPYFRMTRDVAPRIGYKPAL	158

RESULT 14  
AAB96409  
ID AAB96409 standard; Protein; 385 AA

AC	AAB96409;
XX	
DT	29-OCT-2001 (first entry)

DE Putative *P. abyssi* tryptophanyl-tRNA synthetase.  
XX  
KW Hyperthermophilic archaeon, hyperthermophilic protein  
XX  
OS *Pyrococcus abyssi*.

XX	FR2792651-A1.
PN	
XX	
PD	27-OCT-2000.
XX	
PE	21-APR-1999; 99FR-0005034.
XX	
XX	21-APR-1999; 99FR-0005034.
XX	
PA	(CNRS ) CNRS CENT NAT RECH SCI.
PA	(IFRE-) IFREMER INST FR RECH EXPL MER

PI Fortee P, Thierry JC, Prieur D, Dietrich J, Lecompte O;  
PI Querellou J, Weissenbach J, Saurin W, Hellig R;  
DR WPI: 2001-126236/14.  
XX  
XX  
XX New nucleotide sequences isolated from *Pyrococcus abyssi* encode  
PT proteins useful in industry.  
PT

XX  
PS Claim 7; Pages 1099-1100; 1657pp; French.

XX The present invention relates to the genomic sequence of *Pyrococcus*

PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142350.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.

PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 05-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147392.  
PR 06-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148365.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 22-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.







```
1
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162142.
```

```
Query Match 51.2%; Score 1304; DB 21; Length 402;
Best Local Similarity 60.3%; Pred. No. 1.7e-121;
Matches 240; Conservative 67; Mismatches 87; Indels 4; Gaps 2;
```

```
OY 76 DATEAE--EDFVDPMTVOTSSAKGIDYDKLIYRFGSSKIDKELIRIERATGQRPHELR 133
DB 7 DEREAESSQVAVNPWEVSARKDGKIDYDKLIDKFCORLDESLIDRVOLTRSRQPHVFLR 66
OY 134 RGFPSHRDMNOVLDAENKRFELYTGKPSSEAMHGHLPFIETKWLQOVENVPLVI 193
DB 67 RSVFAHRDPNFIIDAYEKGDFLYITGRGPSEALHGLHLPFETIYLQAFKRVPLVI 126
OY 194 QMTDEKYLKMDLTDQAYGDAVENAKDIIACGFDINTFTIFSDDYMGSSGFYKNVVK 253
DB 127 QLTDEKSIWKMLSVESQRLARENAKDIACGFDVTKTFIFRSDPDYVG--GAFYKNVVK 184
OY 254 IOKHTEPNQVKIRFGFTSDCGKISFPAIQAPSFNSFQIIFDRDIDQILICAIQD 313
DB 185 VKCYTLNKAMKIFGFGSDPIAKISFPVQAVPSFSSFPHLFPGKDNRLICAIQD 244
OY 314 DPFYFMTDVAIRIGYPRKALHSTFPFALOGAQTKMSADPNSISFILTAKOIKTKYVN 373
DB 245 DPFYFMTDVAIRIGYPRKALHSTFPFALOGAQTKMSADPNSISFILTAKOIKTKYVN 304
OY 374 KHAESGSGDTJEHRQFSGNCDVVSFMYLTFELEDDEKLEQIRKDYTSGAMLTGELKA 433
DB 305 RYAFSGGQDSIEKHRELKANLEVIDIPVKYLSFLEDDSELEHRIKKEYGEGRMULTGEVKKR 364
```

```
OY 434 LIEVLQPLIAEHQARRKEVTEIVEKFMTPRKLSDFQ 471
DB 365 LIEVTEIVEKRRARAAVTDEMDVAFMAVRPLPSKFE 402

RESULT 10
AAG23697
ID AAG23697 standard; Protein: 426 AA.
AC AAG23697;
DT 17-OCT-2000 (first entry)
DE Arabidopsis thaliana protein fragment SEQ ID NO: 27100.
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hydridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX Arabidopsis thaliana.
OS
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123160.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128274.
XX 16-APR-1999; 99US-0129845.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130449.
XX 23-APR-1999; 99US-0130510.
XX 23-APR-1999; 99US-0130891.
XX 28-APR-1999; 99US-0131449.
XX 30-APR-1999; 99US-0132048.
XX 30-APR-1999; 99US-0132407.
XX 04-MAY-1999; 99US-0132484.
XX 05-MAY-1999; 99US-0132485.
XX 06-MAY-1999; 99US-0132486.
XX 06-MAY-1999; 99US-0132487.
XX 07-MAY-1999; 99US-0132863.
XX 11-MAY-1999; 99US-0134256.
XX 14-MAY-1999; 99US-0134218.
XX 14-MAY-1999; 99US-0134219.
XX 14-MAY-1999; 99US-0134221.
XX 14-MAY-1999; 99US-0134370.
XX 18-MAY-1999; 99US-0134768.
XX 19-MAY-1999; 99US-0134941.
XX 20-MAY-1999; 99US-0135124.
XX 21-MAY-1999; 99US-0135353.
XX 24-MAY-1999; 99US-0135629.
XX 25-MAY-1999; 99US-0136021.
XX 27-MAY-1999; 99US-0136392.
XX 28-MAY-1999; 99US-0136782.
XX 01-JUN-1999; 99US-0137222.
XX 03-JUN-1999; 99US-0137528.
XX 04-JUN-1999; 99US-0137502.
XX 07-JUN-1999; 99US-0137724.
XX 08-JUN-1999; 99US-0138094.
XX 10-JUN-1999; 99US-0138540.
XX 10-JUN-1999; 99US-0138847.
XX 14-JUN-1999; 99US-0139119.
XX 16-JUN-1999; 99US-0139452.
XX 16-JUN-1999; 99US-0139453.
XX 17-JUN-1999; 99US-0139492.
```

KW termination sequence.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EPI033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123160.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 14-JUN-1999; 99US-0138847.  
PR 16-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139452.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141682.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143642.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144684.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 05-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 06-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149920.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151080.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151338.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.





XX The sequences given in AAB47615-18 show full length and truncated  
 CC versions of triphophanyl-tRNA synthetase (TrpRS). The truncated TrpRS of  
 CC the invention comprises a Rossmann fold nucleotide binding domain, and  
 CC is capable of regulating vascular endothelial cell function. It is of  
 CC approx. 40 kDa molecular weight and is produced by cleavage of  
 CC full length TrpRS with polymorphonuclear leucocyte elastase. Truncated  
 CC TrpRS is useful for regulating angiogenesis, tumor metastasis,  
 CC enhancing angiogenesis to a graft, treating myocardial infarction,  
 CC solid tumor, and a condition that would benefit from increased or  
 CC decreased angiogenesis in a mammal, in particular humans. It is also  
 CC useful in diagnosis and as a wound healing agent for treating wounds  
 CC such as dermal ulcers, diabetic ulcers, burns and injuries and in  
 CC plastic surgery when reconstruction is required following a burn or  
 CC for cosmetic purposes. It is particularly useful in the treatment of  
 CC abdominal wounds where there is high risk of infection. Truncated TrpRS  
 CC promotes endothelialization in vascular graft surgery and is used in  
 CC conjunction with angiography to administer the angiogenic tRNA  
 CC synthetase polypeptides or polynucleotides directly to the lumen and  
 CC wall of the blood vessel.

XX  
 SO Sequence 437 AA;

Query Match 91.2%; Score 2324; DB 22; Length 437;  
 Best Local Similarity 100.0%; Pred. No. 1.6e-223;  
 Matches 437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

48 MSYAAAGEDKKACPPENPAPTSNHPDPAEEDFDPWYVOTSSAKGIDYDKLYRF 107  
 Db 1 MSYAAAGEDKKACPPENPAPTSNHPDPAEEDFDPWYVOTSSAKGIDYDKLYRF 60  
 QY 108 GSSKIDKELINRIERATGQRPHHFLRGIFFSHRDMNOVLDAVENKKPFYLYTGRPSSE.167  
 Db 61 GSSKIDKELINRIERATGQRPHHFLRGIFFSHRDMNOVLDAVENKKPFYLYTGRPSSE 120  
 QY 168 AMHGHILPFTFKMLQDVENVPLVITOMTDEKYLKDLTDQAYGDAVENAKDIIACGF 227  
 Db 121 AMHGHILPFTFKMLQDVENVPLVITOMTDEKYLKDLTDQAYGDAVENAKDIIACGF 180  
 QY 228 DINKTIFISDDYMGMSGFFKNVYKIQKHTFNQVKGIFGFTSDSCIGKISFPAIQAP 287  
 Db 181 DINKTIFISDDYMGMSGFFKNVYKIQKHTFNQVKGIFGFTSDSCIGKISFPAIQAP 240  
 QY 288 SFSNSFPQIFRDRDIOCLICPAIDQDYPFRMTDRVAPRIGYPRKALHSTFFPALQGAQ 347  
 Db 241 SFSNSFPQIFRDRDIOCLICPAIDQDYPFRMTDRVAPRIGYPRKALHSTFFPALQGAQ 300  
 QY 348 TKMSASDPNSSIFLDTAKQIKTKVKNKHAISGGRDTEENHQFGNCDVDSFMTLTFEL 407  
 Db 301 TKMSASDPNSSIFLDTAKQIKTKVKNKHAISGGRDTEENHQFGNCDVDSFMTLTFEL 360  
 QY 408 EDDKLEQIRKDYTSNGAMLTEGELKALIEVLOPLAEHQARKKEVTDEIVKPEMTPRKL 467  
 Db 361 EDDKLEQIRKDYTSNGAMLTEGELKALIEVLOPLAEHQARKKEVTDEIVKPEMTPRKL 420  
 QY 468 FDFQKLAALAEHNNHNN 484  
 Db 421 FDFQKLAALAEHNNHNN 437

RESULT 5  
 AAB47617  
 ID AAB47617 standard; Protein: 415 AA.  
 AC AAB47617:  
 XX  
 DT 07-JAN-2002 (first entry)  
 XX  
 DE Human supermini TrpRS.  
 XX  
 KM Tyrosyl-tRNA synthetase; TyRS; Rossmann fold nucleotide binding domain;  
 KM vascular endothelial cell function; burn; plastic surgery; abdomen;  
 KM polymorphonuclear leucocyte elastase; angiogenesis; tumor metastasis;

KW angiogenesis; graft; myocardial infarction; solid tumour; wound healing;  
 KW dermal ulcer; diabetic ulcer; endothelialization;  
 KW triphophanyl-tRNA synthetase; trpRS; vascular graft surgery.  
 OS Homo sapiens.  
 XX  
 XX WO200174841-A1.  
 PD 11-OCT-2001.  
 PF 21-MAR-2001; 2001WO-US08966.  
 PR 31-MAR-2000; 2000US-193471P.  
 PA (SCRI ) SCRI PPS RES INST.  
 PI Schimmel P, Wakasugi K;  
 XX  
 DR WPI; 2001-626377/72.  
 DR N-PSDB; AAA43604.  
 PT New human truncated tyrosyl-tRNA synthetase polypeptide for regulating  
 PT vascular endothelial function, in particular for regulating  
 PT angiogenesis, tumor metastasis and treating myocardial infarction  
 PS Disclosure; Page 129-30; 150pp; English.

XX  
 CC The sequences given in AAB47615-18 show full length and truncated  
 CC versions of triphophanyl-tRNA synthetase (TrpRS). The truncated TrpRS of  
 CC the invention comprises a Rossmann fold nucleotide binding domain, and  
 CC is capable of regulating vascular endothelial cell function. It is of  
 CC approx. 40 kDa molecular weight and is produced by cleavage of  
 CC full length TrpRS with polymorphonuclear leucocyte elastase. Truncated  
 CC TrpRS is useful for regulating angiogenesis, tumor metastasis,  
 CC enhancing angiogenesis to a graft, treating myocardial infarction,  
 CC solid tumor, and a condition that would benefit from increased or  
 CC decreased angiogenesis in a mammal, in particular humans. It is also  
 CC useful in diagnosis and as a wound healing agent for treating wounds  
 CC such as dermal ulcers, diabetic ulcers, burns and injuries and in  
 CC plastic surgery when reconstruction is required following a burn or  
 CC for cosmetic purposes. It is particularly useful in the treatment of  
 CC abdominal wounds where there is high risk of infection. Truncated TrpRS  
 CC promotes endothelialization in vascular graft surgery and is used in  
 CC conjunction with angiography to administer the angiogenic tRNA  
 CC synthetase polypeptides or polynucleotides directly to the lumen and  
 CC wall of the blood vessel.

XX  
 SO Sequence 415 AA;

Query Match 86.1%; Score 2194; DB 22; Length 415;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-210;  
 Matches 414; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

71 SNHGPDAEAEEDFDPWYVOTSSAKGIDYDKLIVRGSSKIDKELINRIERATGQRPH 130  
 Db 2 SNHGPDAEAEEDFDPWYVOTSSAKGIDYDKLIVRGSSKIDKELINRIERATGQRPH 61  
 QY 131 FLRGIFFSHRDMNOVLDAVENKKPFYLYTGRGSSSEAMHGHILPFTFKMLQDVENV 190  
 Db 62 FLRGIFFSHRDMNOVLDAVENKKPFYLYTGRGSSSEAMHGHILPFTFKMLQDVENV 121  
 QY 191 LVIGMTDEKYLKDLTDQAYGDAVENAKDIIACGFDINKTIFISDLDYMGMSGFFKN 250  
 Db 122 LVIGMTDEKYLKDLTDQAYGDAVENAKDIIACGFDINKTIFISDLDYMGMSGFFKN 181  
 QY 251 VVKIQKHTFNQVKGIFGFTSDSCIGKISFPAIQAAFSNSFPQIRDRDIOCLIPCA 310  
 Db 182 VVKIQKHTFNQVKGIFGFTSDSCIGKISFPAIQAAFSNSFPQIRDRDIOCLIPCA 241  
 QY 311 IDQDPFRMTDRVAPRIGYPRKALHSTFFPALQGAOTKMSASDPNSSIFLDTAKQIKT 370  
 Db 242 IDQDPFRMTDRVAPRIGYPRKALHSTFFPALQGAOTKMSASDPNSSIFLDTAKQIKT 301

QY 181 KWLQDVNPLVLTQMTDEKYLKMDLTLDQAYGDAVENADIIACGFDINKTFISDLDY 240  
 DB 185 KWLQDVNPLVLTQMTDEKYLKMDLTLDQAYGDAVENADIIACGFDINKTFISDLDY 244  
 QY 241 MGNSSGYYKNNVVKIQKHVTENQVKGIFGFTDSDICGKISFPALQAAFSNSFPQIFRDR 300  
 DB 245 MGNSSGYYKNNVVKIQKHVTENQVKGIFGFTDSDICGKISFPALQAAFSNSFPQIFRDR 304  
 QY 301 TDIOCLIPCALIDDDPYFRMRDVAAPRIGYPRKPALHSTFFPALQAGATKMSASDPNSSIF 360  
 DB 305 TDIOCLIPCALIDDDPYFRMRDVAAPRIGYPRKPALHSTFFPALQAGATKMSASDPNSSIF 364  
 QY 361 LTTAKOIKTKVKNKHAFFSGGRDTIEEHROFGNCDVVSFMYLTFFLEDDDKLEQIRKDY 420  
 DB 365 LTTAKOIKTKVKNKHAFFSGGRDTIEEHROFGNCDVVSFMYLTFFLEDDDKLEQIRKDY 424  
 QY 421 TSGAMLTGELKALIEVLQPLIAEHQARREVTDEIVKEFMTPRKLSFDRQ 471  
 DB 425 TSGAMLTGELKALIEVLQPLIAEHQARREVTDEIVKEFMTPRKLSFDRQ 475

## RESULT 3

AA05372  
 ID AAY05372 standard; Protein: 471 AA.

AC AAY05372;  
 DT 30-JUN-1999 (first entry)

DE Human HCMV inducible gene protein, SEQ ID NO 12.

KW HCMV inducible gene; c1g; human; human cytomegalovirus; interferon;  
 KM anti-viral therapy; anti-HCMV therapy; detection; diagnosis;  
 XX drug screening.

OS Homo sapiens.

PN WO9913075-A2.

PD 18-MAR-1999.

PF 08-SEP-1998; 98WO-US18638.

PR 22-SEP-1997; 97OS-0059725.  
 PR 08-SEP-1997; 97US-0058180.

PA (UYPR-) UNITV PRINCETON.

PI Cong J, Schenk T, Zhu H;

DR WPI: 1999-243729/20.

DR N-PSDB: AAX33942.

PT New isolated human genes

PS Claim 3; Page 112-114; 184pp; English.

XX This sequence is encoded by a human gene of the invention, and is induced  
 CC to express by both HCMV and interferon (IFN), designated HCMV-inducible  
 CC genes (c1g or c1gs). The invention also relates to genes that are  
 CC repressed in the presence of HCMV infection, designated HCMV-repressible  
 CC genes (crg or crgs). The products can be used to obtain agents which can  
 CC be used for anti-viral therapy, particularly anti-HCMV therapy. They can  
 CC also be used for the development of drugs that would allow for higher  
 CC dosage IFN treatments without the concomitant toxicity normally  
 CC associated with administering high levels of IFN. The products can also  
 CC be used for detection, diagnosis and drug screening.

XX Sequence 471 AA;

Query Match 96.2%; Score 2449; DB 20; Length 471;

Best Local Similarity 99.4%; Pred. No. 5.7e-236;  
 Matches 468; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MPNSEPASLLELFNSIATOGELVSLKAGNASKDEIDSAVKMLVSLKMSYKAAAGEDYKA 60  
 DB 1 MPNSEPASLLELFNSIATOGELVSLKAGNASKDEIDSAVKMLVSLKMSYKAAAGEDYKA 60  
 QY 61 DCPGPNAPPTSNMGPDATAEEDFVDPWTVOTSSAKGIDYDKLIVRGSSKIDKELINRI 120  
 DB 61 DCPGPNAPPTSNMGPDATAEEDFVDPWTVOTSSAKGIDYDKLIVRGSSKIDKELINRI 120  
 QY 121 ERATGQRPHFLRGRIFESHRDMQVLDAYENKRPFLYGRGSSAMHVGHLIPFT 180  
 DB 121 ERATGQRPHFLRGRIFESHRDMQVLDAYENKRPFLYGRGSSAMHVGHLIPFT 180  
 QY 181 KWLQDVNPLVLTQMTDEKYLKMDLTLDQAYGDAVENADIIACGFDINKTFISDLDY 240  
 DB 181 KWLQDVNPLVLTQMTDEKYLKMDLTLDQAYGDAVENADIIACGFDINKTFISDLDY 240  
 QY 241 MGNSSGYYKNNVVKIQKHVTENQVKGIFGFTDSDICGKISFPALQAAFSNSFPQIFRDR 300  
 DB 241 MGNSSGYYKNNVVKIQKHVTENQVKGIFGFTDSDICGKISFPALQAAFSNSFPQIFRDR 300  
 QY 301 TDIOCLIPCALIDDDPYFRMRDVAAPRIGYPRKPALHSTFFPALQAGATKMSASDPNSSIF 360  
 DB 301 TDIOCLIPCALIDDDPYFRMRDVAAPRIGYPRKPALHSTFFPALQAGATKMSASDPNSSIF 360  
 QY 361 LTTAKOIKTKVKNKHAFFSGGRDTIEEHROFGNCDVVSFMYLTFFLEDDDKLEQIRKDY 420  
 DB 361 LTTAKOIKTKVKNKHAFFSGGRDTIEEHROFGNCDVVSFMYLTFFLEDDDKLEQIRKDY 420  
 QY 421 TSGAMLTGELKALIEVLQPLIAEHQARREVTDEIVKEFMTPRKLSFDRQ 471  
 DB 421 TSGAMLTGELKALIEVLQPLIAEHQARREVTDEIVKEFMTPRKLSFDRQ 471

## RESULT 4

AA047616  
 ID AAB47616 standard; Protein: 437 AA.

AC AAB47616;

DT 07-JAN-2002 (first entry)

DE Human mini T1PRs.

KW Tyrosyl-tRNA synthetase; TYPRS; Roessmann fold nucleotide binding domain;  
 KM vascular endothelial cell function; burn; plastic surgery; abdomen;

KW polymorphonuclear leucocyte elastase; angiogenesis; tumour metastasis;  
 KM dermal ulcer; diabetic ulcer; endohehialization; solid tumour; wound healing;

KW tryptophanyl-tRNA synthetase; trprs; vascular graft surgery.

XX Homo sapiens.

PN WO200174841-A1.

PD 11-OCT-2001.

PF 21-MAR-2001; 2001WO-US08966.

PR 31-MAR-2000; 2000US-193471P.

PA (SCRI) SCRIPPS RES INST.

PI Schimmel P, Makasugi K;

DR WPI: 2001-626377/72.

DR N-PSDB: AAH43603.

PT New human truncated tyrosyl-tRNA synthetase polypeptide for regulating  
 PT vascular endothelial function, in particular for regulating

PT angiogenesis, tumor metastasis and treating myocardial infarction -

PS Disclosure; Page 123-24; 150pp; English.



PT New human truncated tyrosyl-tRNA synthetase polypeptide for regulating  
PT vascular endothelial function, in particular for regulating  
XX angiogenesis, tumor metastasis and treating myocardial infarction  
XX  
PS Disclosure; Page 117-19; 150pp; English.  
XX  
CC The sequences given in AAB47615-18 show full length and truncated  
CC versions of triphosphanyl-tRNA synthetase (TrpRS). The truncated TrpRS of  
CC the invention comprises a Rossmann fold nucleotide binding domain, and  
CC is capable of regulating vascular endothelial cell function. It is of  
CC approx. 40 kilo Dalton molecular weight and is produced by cleavage of  
CC full length TrpRS with polynucleonuclear leucocyte elastase. Truncated  
CC TrpRS is useful for regulating angiogenesis, tumor metastasis,  
CC enhancing angiogenesis to a graft, treating myocardial infarction,  
CC solid tumor, and a condition that would benefit from increased or  
CC decreased angiogenesis in a mammal, in particular humans. It is also  
CC useful in diagnosis and as a wound healing agent for treating wounds  
CC such as dermal ulcers, diabetic ulcers, burns and injuries and in  
CC plastic surgery when reconstruction is required following a burn or  
CC for cosmetic purposes. It is particularly useful in the treatment of  
CC abdominal wounds where there is high risk of infection. Truncated TrpRS  
CC promotes endothelialization in vascular graft surgery and is used in  
CC conjunction with angiography to administer the angiogenic tRNA  
CC synthetase polypeptides or polynucleotides directly to the lumen and  
CC wall of the blood vessel.  
XX  
XX

SQ Sequence 484 AA;

Query Match 100.0%; Score 2547; DB 22; Length 484;

Best Local Similarity 100.0%; Pred. No. 9.2e-246;

Matches 484; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPNSEPASLLEFNSIATOGELVRSIKAGNASKDEIDSAVKMLVSLKMSYKAAAGEDYKA 60  
DB 1 MPNSEPASLLEFNSIATOGELVRSIKAGNASKDEIDSAVKMLVSLKMSYKAAAGEDYKA 60  
QY 61 DCPGPNAPPTNSNGPDATAEEDFVDPWTYOTSSAKGIDYDKLIVRGSSKIDKELINRI 120  
DB 61 DCPGPNAPPTNSNGPDATAEEDFVDPWTYOTSSAKGIDYDKLIVRGSSKIDKELINRI 120  
QY 121 ERATGGRPHHFLRGIFFSHRDMNOVLDAVENKKPFYLYTGRGSSSEAMVGHILPIFT 180  
DB 121 ERATGGRPHHFLRGIFFSHRDMNOVLDAVENKKPFYLYTGRGSSSEAMVGHILPIFT 180  
QY 181 KWLQDFVNPVLVQMTDDEYLYMKDLTLDOAYGDVAENANDIACGFDINKTFEISDL 240  
DB 181 KWLQDFVNPVLVQMTDDEYLYMKDLTLDOAYGDVAENANDIACGFDINKTFEISDL 240  
QY 241 MGSMSGFYKVVVIOKHTFENOVYKGFEGFTDSCIGKISFPALQAAFSFNSFPQIFRDR 300  
DB 241 MGSMSGFYKVVVIOKHTFENOVYKGFEGFTDSCIGKISFPALQAAFSFNSFPQIFRDR 300  
QY 301 TDIOCLIPCAIDDDPYFRMTDVAAPRIGYKPKPALHSTFPALQGAQTAKSADPNSSIF 360  
DB 301 TDIOCLIPCAIDDDPYFRMTDVAAPRIGYKPKPALHSTFPALQGAQTAKSADPNSSIF 360  
QY 361 LDTAKOIKTKVKNHAFSGGRDTIEEHROFGNCDVDSFMYLTFFLEDDDKLEQIRKDY 420  
DB 361 LDTAKOIKTKVKNHAFSGGRDTIEEHROFGNCDVDSFMYLTFFLEDDDKLEQIRKDY 420  
QY 421 TSCAMLTGELKKALIEVLOPLIAHOARRKEVDETVKEEMPRKLSFPOKLAALLENH 480  
DB 421 TSCAMLTGELKKALIEVLOPLIAHOARRKEVDETVKEEMPRKLSFPOKLAALLENH 480  
QY 481 HHHH 484  
DB 481 HHHH 484

RESULT 2  
AAB58220  
ID AAB58220 standard; Protein; 475 AA.  
XX

AC AAB58220;  
XX  
DT 14-MAR-2001 (first entry)  
XX  
DE Lung cancer associated polypeptide sequence SEQ ID 558.

KW Human; lung cancer associated protein; neuroprotective; cytostatic;  
KW cardioactive; immunomodulatory; muscular active; vulnary;  
KW gastrointestinal; nephrotropic; antineoplastic; gynecological;  
KW antibacterial; diagnosis; neural disorder; immune disorder; reproductive;  
KW proliferative disorder; wound healing; infectious disease.

OS Homo sapiens.

PN MO200055180-A2.

PD 21-SEP-2000.

PF 08-MAR-2000; 2000MO-0505918.

PR 12-MAR-1999; 99US-0124270.

PA (HUMA-) HUMAN GENOME SCI INC.  
(ROSE/) ROSEN C A.

PI Ruben SM;

DR WPI: 2000-587514/55.

DR N-PSDB; AAF18096.

PT Lung cancer associated gene sequences, referred to as lung cancer  
PT antigens, useful for treatment, prevention, and diagnosis of disorders  
PT such as lung cancer .  
XX  
XX

PS Claim 11; Page 1052-1053; 1425pp; English.

CC Polynucleotide sequences AAF17982 - AAF18424 encode human lung cancer  
CC associated proteins represented in AAB58106 - AAB58548. Lung cancer  
CC associated proteins and polynucleotide sequences, their agonists, and  
CC antagonists may have neuroprotective; cytostatic; cardioactive;  
CC immunomodulatory; muscular active general; vulnary; gastrointestinal  
CC general; nephrotropic; antineoplastic; gynecological; or antibacterial  
CC activity. The invention also includes antibodies specific for the  
CC protein or polynucleotide sequences. The lung cancer associated  
CC polynucleotide sequences may be used for detection of lung cancer,  
CC chromosome identification, as chromosome markers, and for numerous other  
CC diagnostic or research purposes. The proteins may be used to treat  
CC disorders such as neural, immune, muscular, reproductive,  
CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative  
CC disorders. The proteins may also be used in the treatment of wounds and  
CC infectious diseases. Polynucleotide sequences AAF18425 - AAF18433 and  
CC peptide AAB58549 are used in the course of the invention for the  
CC identification and characterisation of the polynucleotide and protein  
CC sequences.  
XX  
XX

SQ Sequence 475 AA;

Query Match 96.3%; Score 2454; DB 21; Length 475;

Best Local Similarity 99.6%; Pred. No. 1.8e-236;

Matches 469; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MPNSEPASLLEFNSIATOGELVRSIKAGNASKDEIDSAVKMLVSLKMSYKAAAGEDYKA 60  
DB 1 MPNSEPASLLEFNSIATOGELVRSIKAGNASKDEIDSAVKMLVSLKMSYKAAAGEDYKA 60  
QY 5 MPNSEPASLLEFNSIATOGELVRSIKAGNASKDEIDSAVKMLVSLKMSYKAAAGEDYKA 64  
DB 5 MPNSEPASLLEFNSIATOGELVRSIKAGNASKDEIDSAVKMLVSLKMSYKAAAGEDYKA 64  
QY 61 DCPGPNAPPTNSNGPDATAEEDFVDPWTYOTSSAKGIDYDKLIVRGSSKIDKELINRI 120  
DB 61 DCPGPNAPPTNSNGPDATAEEDFVDPWTYOTSSAKGIDYDKLIVRGSSKIDKELINRI 124  
QY 121 ERATGGRPHHFLRGIFFSHRDMNOVLDAVENKKPFYLYTGRGSSSEAMVGHILPIFT 180  
DB 121 ERATGGRPHHFLRGIFFSHRDMNOVLDAVENKKPFYLYTGRGSSSEAMVGHILPIFT 184

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 24, 2002, 12:49:40 : Search time 30.3193 seconds  
(without alignments)  
1773.117 Million cell updates/sec

Title: US-09-813-718-10

Perfect score: 2547  
Sequence: 1 MPNEPASILFELFNSTATQG.....KLSPDFOKLAALHHHHHHH 484

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : A.Geneseq\_032802:\*

1: /SIDSI/gcgdata/geneseq/geneseqp-emb1/1  
2: /SIDSI/gcgdata/geneseq/geneseqp-emb1/1  
3: /SIDSI/gcgdata/geneseq/geneseqp-emb1/1  
4: /SIDSI/gcgdata/geneseq/geneseqp-emb1/1  
5: /SIDSI/gcgdata/geneseq/geneseqp-emb1/1  
6: /SIDSI/gcgdata/geneseq/geneseqp-emb1/1  
7: /SIDSI/gcgdata/geneseq/geneseqp-emb1/1  
8: /SIDSI/gcgdata/geneseq/geneseqp-emb1/1  
9: /SIDSI/gcgdata/geneseq/geneseqp-emb1/1  
10: /SIDSI/gcgdata/geneseq/geneseqp-emb1/1  
11: /SIDSI/gcgdata/geneseq/geneseqp-emb1/1  
12: /SIDSI/gcgdata/geneseq/geneseqp-emb1/1  
13: /SIDSI/gcgdata/geneseq/geneseqp-emb1/1  
14: /SIDSI/gcgdata/geneseq/geneseqp-emb1/1  
15: /SIDSI/gcgdata/geneseq/geneseqp-emb1/1  
16: /SIDSI/gcgdata/geneseq/geneseqp-emb1/1  
17: /SIDSI/gcgdata/geneseq/geneseqp-emb1/1  
18: /SIDSI/gcgdata/geneseq/geneseqp-emb1/1  
19: /SIDSI/gcgdata/geneseq/geneseqp-emb1/1  
20: /SIDSI/gcgdata/geneseq/geneseqp-emb1/1  
21: /SIDSI/gcgdata/geneseq/geneseqp-emb1/1  
22: /SIDSI/gcgdata/geneseq/geneseqp-emb1/1

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2547	100.0	484	22	AAB47615 Human full-length
2	2454	96.3	475	21	AAB58220 Lung cancer associ
3	2449	96.2	471	20	AAY05372 Human HCMV induci
4	2324	91.2	437	22	AAB47616 Human mini TyRS.
5	2194	86.1	415	22	AAB47617 Human supermini tr
6	2066	81.1	392	22	AAB47618 Human inactive Tyr
7	1378	54.1	430	22	ABB64621 Drosophila melanog
8	1378	54.1	430	22	ABB67203 Arabidopsis thalia
9	1304	51.2	402	21	AAG23698 Arabidopsis thalia
10	1304	51.2	426	21	AAG23697 Arabidopsis thalia
11	1218.5	47.8	424	22	AAB66931 Tyrophenyl-L-TRNA

12	968	38.0	292	21	AAG23699 Arabidopsis thalia
13	831	32.6	173	21	AAB58517 Lung cancer associ
14	803	31.5	385	22	AAB96409 Putative P. abyss
15	579.5	22.8	142	21	AAB58219 Lung cancer associ
16	269	10.6	71	22	AAO10441 Human polyprotein
17	185.5	7.3	341	19	AAW56423 Tyrophenyl-L-TRNA
18	185.5	7.3	341	22	AAU37851 Streptococcus pneu
19	173	6.8	372	22	AAB47614 Human mini TyRS.
20	171.5	6.7	372	22	AAB47612 Human mini TyRS.
21	169.5	6.7	409	20	AAW94248 C. albicans tyrosy
22	167.5	6.6	409	20	AAW94247 C. albicans tyrosy
23	159.5	6.3	379	20	AAV37623 Protein involved i
24	158.5	6.2	378	22	AAW96600 Putative P. abyss
25	155	6.1	1714	22	ABW61487 Drosophila melanog
26	153	6.0	1512	22	AAW78732 Human protein seq
27	153	6.0	1512	22	AAW04349 Mammalian toxicolo
28	153	6.0	1550	22	AAW79716 Human protein seq
29	152	6.0	344	20	AAV35439 Chlamydia pneumoni
30	146	5.7	339	22	AAU35829 Helicobacter pylor
31	144	5.7	525	22	ABW60745 Arabidopsis thalia
32	138.5	5.4	140	21	AAW23032 Arabidopsis thalia
33	137	5.4	536	22	AAB47611 Human full length
34	128	5.0	334	22	AAU35477 Haemophilus influe
35	123	4.8	347	21	AAW31904 Arabidopsis thalia
36	122.5	4.8	346	21	AAW48617 Arabidopsis thalia
37	120.5	4.7	404	22	AAU54020 Propionibacterium
38	115.5	4.5	129	21	AAW23033 Arabidopsis thalia
39	113.5	4.5	418	22	AAE02578 Enterococcus tyros
40	113.5	4.5	423	22	AAU35085 Enterococcus faeca
41	113	4.4	415	22	AAU33998 Staphylococcus aur
42	113	4.4	420	18	AAW19781 Tyrosyl-L-TRNA synth
43	113	4.4	420	22	AAU37056 Staphylococcus aur
44	112	4.4	345	22	AAW90498 C glutamicum prote
45	108	4.2	334	22	AAU38213 Salmonella typhi c

# ALIGNMENTS

RESULT 1	AAW47615	standard; Protein; 484 AA.
ID	AAW47615	standard; Protein; 484 AA.
XX	AAW47615:	
AC	07-JAN-2002	(first entry)
XX		
DT		
XX		
XX		
DE		Human full-length TyRS.
XX		
XX		Tyrosyl-L-TRNA synthetase; TyRS; Rossmann fold nucleotide binding domain;
KW		vascular endothelial cell function; burn; plastic surgery; abdomen;
KW		polymorphonuclear leucocyte elastase; angiogenesis; tumour metastasis;
KW		angiogenesis; graft; myocardial infarction; solid tumour; wound healing;
KW		dermal ulcer; diabetic ulcer; endothelialization;
KW		tyrophenyl-L-TRNA synthetase; tyRS; vascular graft surgery.
OS		Homo sapiens.
XX		
PN	WO200174841-A1.	
XX		
PD	11-OCT-2001.	
XX		
XX		
PF	21-MAR-2001; 2001WO-US08966.	
XX		
PR	31-MAR-2000; 2000US-193471P.	
XX		
PA	(SCRI ) SCRIPPS RES INST.	
XX		
PI	Schimmel P, Wakasugi K.	
XX		
DR	WPI: 2001-626377/72.	
XX		
DR	N-PSDB: AAH43602.	
XX		

[illegible]

RESULT 15		
Q9HN66		
ID	Q9HN66	PRELIMINARY; PRT; 380 AA.
AC	Q9HN66;	
DT	01-MAR-2001 (TREMBLrel. 16, Created)	
DT	01-MAR-2001 (TREMBLrel. 16, Last sequence update)	
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)	
DE	TRYPTOPHANYL-TRNA SYNTHETASE.	
OS	TRPS2 OR VNG3232G.	
SN	Halobacterium sp. (strain NRC-1).	
OC	Archaea; Euryarchaeota; Halobacteriales; Halobacteriaceae;	
OC	Halobacterium.	
OX	NCBI_TaxID=64091;	

RP SEQUENCE FROM N.A.  
RX MEDLINE=20504483; Pubmed=11016950;  
RA N.W.V., Kennedy S.P., Mahairas G.C., Bergquist B., Pan M.,  
RA Shukla H.D., Lasky S.R., Baliga N.S., Thorsson V., Sbrogna J.,  
RA Swartzell S., Weir D., Hall J., Dahl T.A., Welt R., Goo Y.A.,  
RA Ledthauser B., Keller K., Cruz R., Danson M.J., Hough D.W.,  
RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angelina C.M., Dale H.,  
RA Isenberger T.A., Peck R.F., Pohlschroder M., Spidlich J.L., Jung K.-H.,  
RA Alam M., Fretels T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,  
RA Edhardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.;  
RT "Genome sequence of *Halobacterium* species NRC-1.";  
RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).  
DR EML: AEO05109: AAG20355.1; -;  
DR InterPro: IPR001994: Cytidylyltransferase.  
DR InterPro: IPR002305: tRNA-synt.1b.  
DR InterPro: IPR001412: tRNA-synt.1.  
DR InterPro: IPR002306: tRNA-synt.tlp.  
DR Pfam: PF00579: tRNA-synt.1b; 1.  
DR PRINTS: PR01039; TRNASYNTHTRP.  
DR PROSITE: PS00178: AA\_TRNA\_LIGASE\_I; UNKNOWN.1.  
KW Aminoacyl-tRNA synthetase; complete proteome.  
SQ SEQUENCE 380 AA; 41936 MW; 0F7B2B953586404F CRC64;

Query Match	25.3%	Score 534.5	DB 17	Length 380
Best Local Similarity	35.4%	Pred. NO. 1.4e-38		
Matches 141; Conservative	76	Mismatches 140	Indels 41	Gaps 17

```

QY 10 AAEDEVDPMTQTSASAKIOXDKI.VLFRGSSKIDKELINIERQTGRPHNLKRGIFPS 69
Db 3 ADGNVTFYAEVSD--LDYKELLRAGADELTDODRAPR-----DPLVNRGLFVA 53
QY 70 HRDMNQVLDAYENKRPFLYLRGSPSSSEAMVGHILPFIFETKWLQDVFNPVLVQMTDDE 129
Db 54 GRVDVDDFLTAGES---YLVGVGPG--PMLHGAMVFYFARRLODFGARVYVPLSDDE 108
QY 130 KYLMKDLTLDAQGYDAVE-NAKDLACGDFINKTFTF---SDDYV-GMSGGYKKNVVKI 184

```

```

Db      109 KYWFDQIPAEI -GDYLKANKRLDILNAGFDEELIRIVDRDADVLXPLATAFAGDV --- 164
OY      185 OKHVTFNOYKGI FGFETSDICIGKISFPALQAPSFNSFOIFDRPTDIOCLIPCAIDQ 244
Db      165 -RHALLQWVYG----EPDNGQAFYPAVQTANHL---LQLVNG--EHEFLVPIADVD 213
OY      245 PYFRRTRDVABRIGP--KPALISTFEPALOGAQTKMSADPNSSIFLDTFKQIKTKV 302
Db      214 PHVRSRGRVAAKARRPYVCKPALLMQFLPISLAG -PGKMSSS-AGVSRILDSPTVAENEV 271
OY      303 NKHAFSGRDITIEEHROFGNCDVDVDSWMLTFPLEDD- KLEQIRKDYVSGAMLTGEL 361
Db      272 RTHAATGGRAVVEEHRAGGVPADVDVFQYLSAFEEPDDEALAIREREYRAGDILSELK 331
OY      362 KALIEVLOPLIAHQARKEVTEDEIVEFMPRLTSP 399
Db      332 DLAADRITEEFLAAHQRRALGCD--VTEALAEFLITDD 367

```

Search completed: October 24, 2002, 12:53:05  
Job time : 22.7502 secs

```

OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FRIEDLIN;
RA Bothe G., Pohl T., Ivens A.C., Lawson D., Quail M., Rajandream M.A.,
RA Barrell B.G.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=FRIEDLIN;
RA MEDLINE=9816435; PubMed=9477341;
RA Ivens A.C., Lewis S.M., Bagherzadeh A., Zhang L., Chan H.M.,
RA Smith D.F.;
RL "A physical map of the Leishmania major Friedlin genome.";
RL Genome Res. 8:135-145(1998).
DR EMBL: AL121862; CAB58393.1;
DR InterPro: IPR002305; tRNA-synt_1b.
DR InterPro: IPR001412; tRNA-synt_1.
DR InterPro: IPR002306; tRNA-synt_trp.
DR Pfam: PF00579; tRNA-synt_1b; 1.
DR PRINTS: PR01039; tRNAsyntHTRP.
DR PROSITE: PS00178; AA_TRNA_LIGASE_I; 1.
DR Aminoacyl-tRNA synthetase.
SQ SEQUENCE 490 AA; 54904 MW; 172c68622c511d3d CRC64;

Query Match 37.6%; Score 795.5; DB 5; Length 490;
Best Local Similarity 38.1%; Pred. No. 2.4e-61;
Matches 173; Conservative 87; Mismatches 129; Indels 65; Gaps 11;

OY 3 HGPDATEEDFDVDPMTVTQTSANGIDYDKLIVRFGSSKID-----KELINRIERAT- 54
DB 40 HPRGDAGADVDITPMVVAKPGQINVDYRLTFKABRMDDGARQHMADVMACRKRMT 99
OY 55 -----GQR-----HFLRIGIFSRDM 73
DB 100 TTPPSANAEVAPFVLDAVPAAGAAIQAASPAHQORQPPALALHFFHFDIAFSHRL 159
OY 74 NOVLDAYE-----NKKPFYLYTGSGPSEAMVHGLIPFTFKWLODFVNPVLIQMTDD 128
DB 160 HKALVLDLEASVKGERSEVFLTGSPSGIMHGLVLPMLTKLQDFSLPLVLIQTDD 219
OY 129 EKYIMKDLTDQAYGAV--ENAKDIIACGFDINKTFIFSDLDGMSSGFGYKNVAK 186
DB 220 EKFLFRDVPPEGAKADBLIRSNKIDIIAFNPNRHTFIFRTHYMG--DMYPTVLROR 276
OY 187 HVPFNQKGIPTGDSICIKISPPALQAPSPNSFPQIFR--DRDIDCLIPCAIDOD 244
DB 277 SMTGNNAVKHLLGITDSNVKLAEPATQAPCFSTAFRVLRLNGDR--PMRCLIPCAIDOD 335
OY 245 PYFRMTDVPAPRIGYPRPALHSTFPALQGAOTKMSAS--DPNSIFLPTAKOIKTKVN 303
DB 336 PFFVLTPTAALRLKQLPALHRTKLPALGLEKSSABEKGVTLHTDKOYRKKL- 394
OY 304 KHAFGSGRDTIEEHROFGNCDVDSFVLTFFLEDDDKLEQIKRDYTSAMLTELKKA 363
DB 395 RRAFSGGCATLEQMGHTGANLDELDAVAYGLRFPCPDPTLRADYQYRRSGTINSGEVKD 454
OY 364 LIE-VIQLPLAEHQARKKEYTDELVEKFTPRKL 396
DB 455 AADCIIREVLHDMRRERATVTDVVEFCRIRDI 488

RESULT 13
OY 095295 PRELIMINARY; PRT; 136 AA.
AC 095295;
DT 01-FEB-1997 (Tremblrel. 02, Created)
DT 01-FEB-1997 (Tremblrel. 02, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE TRYPTOPHANYL-TRNA SYNTHASE (FRAGMENT).
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=SMALL INTESTINE;
RA Winteroe A.K., Fredholm M., Davies W.;
RL "Evaluation and characterization of a porcine small intestine cDNA
RT library.";
RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL: Z81267; CAB03585.1;
DR InterPro: IPR002306; tRNA-synt_trp.
DR PRINTS: PR01039; tRNAsyntHTRP.
FT NON_TER 1
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 136 AA; 15404 MW; 951F75D1B9CD0617 CRC64;

Query Match 32.2%; Score 682; DB 6; Length 136;
Best Local Similarity 94.1%; Pred. No. 3.6e-52;
Matches 128; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

OY 211 PAIOAPSFNSFPQIFRDRDIDCLIPCAIDODPYFRMTDVPAPRIGYPRPALHSTFF 270
DB 1 PAIOXAPSFSSFPQIFRDRDIDCLIPCAIDODPYFRMTDVPAPRIGYPRPALHSTFF 60
OY 271 PALQAGOTKMSASDPNSIFLPTAKOIKTKVNHAFSGGGRDTEEHROFGNCDVDSF 330
DB 61 PALQAGOTKMSAXPNSSIFLPTAKOIKTKVNHAFSGGGRDTEEHROFGNCDVDSF 120
OY 331 MYLTFLEDDDKLEQI 346
DB 121 MYLTFLEDDDKRQEI 136

RESULT 14
OY 059584 PRELIMINARY; PRT; 301 AA.
AC 059584;
DT 01-AUG-1998 (Tremblrel. 07, Created)
DT 01-AUG-1998 (Tremblrel. 07, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE 301AA LONG HYPOTHETICAL TRYPTOPHANYL-TRNA SYNTHETASE.
GN PH1921.
OS Pyrococcus horikoshii.
OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.
OX NCBI_TaxID=53953;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OT3;
RX MEDLINE=98344137; PubMed=9679194;
RA Kawarabayashi Y., Sawada M., Horikawa H., Haikawa Y., Hino Y.,
RA Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,
RA Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,
RA Funahashi T., Tanaka T., Kudoh Y., Yamazaki J., Kuslida N., Oguchi A.,
RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,
RA Masuchi Y., Shizuya H., Kikuchi H.;
RL "Complete sequence and gene organization of the genome of a hyper-
RT thermophilic archaeobacterium, Pyrococcus horikoshii OT3.";
RL DNA Res. 5:55-76(1998).
DR EMBL: AP000007; BAA31046.1;
DR InterPro: IPR002305; tRNA-synt_1b.
DR InterPro: IPR002306; tRNA-synt_trp.
DR Pfam: PF00579; tRNA-synt_1b; 1.
DR PRINTS: PR01039; tRNAsyntHTRP.
KW Aminoacyl-tRNA synthetase; Complete proteome.
SQ SEQUENCE 301 AA; 35312 MW; 957143E1647CAE1 CRC64;

Query Match 29.6%; Score 626.5; DB 17; Length 301;
Best Local Similarity 44.2%; Pred. No. 8.3e-47;
Matches 134; Conservative 56; Mismatches 92; Indels 21; Gaps 7;

OY 99 MHVGHLLPIFTKWLQDFVNPVLIQMTDDEKYLMDK-LTLDQAYGDAVENARDIACGF 157
DB 1 MHIGHIIPFATKWLQDKFGVNIYIQTDEKFLFKENLTFDDTKRWAYDNIIIDIAVGF 60

```

Db 201 IWKVVNTQARAIKFGFTEDDCIGKAAFPAAVEAACPFASSFPQIEFKRNDICLIPCAIDQ 260  
Qy 244 DPEFMTDVAAPRIGYPPKALHSTFPALOGAOTKMSADPNSSIFLTPAKOIKTIV 302  
Db 261 DPEFMTDVAAPRIGYPPKALHSTFPALOGAOTKMSADPNSSIFLTPAKOIKTIV 319

## RESULT 10

Q976M1 PRELIMINARY; PRT; 406 AA.  
AC Q976M1;  
DT 01-DEC-2001 (TREMBLrel. 19, Created)  
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE PUTATIVE TRYPHOPHANYL-TRNA SYNTHETASE.  
GN ST0169.  
OS Sulfolobus tokodaii.  
OC Archaea; Crenarchaeota; Sulfolobales; Sulfolobaceae; Sulfolobus.  
OX NCBI\_TaxID=11955;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-JCM 10545 / 7;  
RX PubMed=11572479;  
RA Kawarabayashi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,  
RA Sekine M., Baba S.-I., Ankaei A., Kosugi H., Hosoyama A., Fukui S.,  
RA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,  
RA Yoshizawa T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,  
RA Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A.,  
RA Oshima T., Kikuchi H.;  
RT "Complete genome sequence of an aerobic thermophilic  
RT Crenarchaeon, Sulfolobus tokodaii strain7.";  
RL DNA Res. 8:123-140(2001).  
KW EMBL: AP000981; BAB5126.1;  
KW Aminoacyl-tRNA synthetase; Hypothetical protein; Complete proteome.  
SQ SEQUENCE 406 AA; 47979 MW; 276CD6506BC94B3 CRC64;

Query Match 43.0%; Score 910.5; DB 17; Length 406;

Best Local Similarity 50.5%; Pred. No. 1,5e-71;  
Matches 189; Conservative 63; Mismatches 99; Indels 23; Gaps 9;

Qy 12 EDFVDPWTQVTSKAG-IDYDKLIVRGSSKIDKELINRIERATGQRPHNLRGIGFS 69  
Db 28 QDFNTPMEV-----KGKVDYDKLIVQFGTKITSELEKIKSIINDELHVMRLRDVFFS 82  
Qy 70 HDMMQVLDAYENKRPFLYITGRGSSSEAMHGHILPFTKWLQDVENVPLVIQMTDE 129  
Db 83 HMDLDLVKLDYDQGGFFLYTGRAPSL-GMHIGHILPFTKWLQDVENVPLVIETDE 141  
Qy 130 KILMK-DLTLDAYGDAVENAKDIIACGFDINKTFISDLDMGSSGFEKYNVVIQKHV 188  
Db 142 KMRNPETLDDQTRQMAYNIIIDIIAVGFNPDKTFIFQDTEYI---RNNYPIAIKAKKL 198  
Qy 189 TENQYKGIFFGTFSDICIGKISFPALOGAOTKMSADPNSSIFLTPAKOIKTIV 248  
Db 199 TSEVATGATGLTSSNIGIYTPALQIAPT-----MEKR---RCLIPAGIDDPYWR 248  
Qy 249 MRDVAAPRIGYPPKALHSTFPALOGAOTKMSADPNSSIFLTPAKOIKTIV 308  
Db 249 LQRDIAESIGYKAAQIHSKFLPPLTGPBGKSSSQPEAIYLVDPKVERKIMKYAFS 308  
Qy 309 GGRDITEEHRQFGNCDVVSFMYLTFLE-DDDKLEQIRKDYSGAMLTGELKALIV 367  
Db 309 GGQPTTELHRRKYGNDIDIVSFWMLYFEPDNKIKKIEEDYRSQALLTGELKOLIEK 368  
Qy 368 LQPLIAEHQARRKE 381  
Db 369 LMDLEEHRRKEE 382

## RESULT 11

Q972X0 PRELIMINARY; PRT; 386 AA.

AC Q972X0;  
DT 01-OCT-2001 (TREMBLrel. 18, Created)  
DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)  
DE 01-OCT-2001 (TREMBLrel. 18, Last annotation update)  
DE TRYPHOPHANYL-TRNA SYNTHETASE (TRPS) (EC 6.1.1.2).  
GN TRPS.  
OS Sulfolobus solfataricus.  
OC Archaea; Crenarchaeota; Sulfolobales; Sulfolobaceae; Sulfolobus.  
OX NCBI\_TaxID=2287;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-ATCC 35092 / DSM 1617 / P2;  
RX MEDLINE=21332296; PubMed=11427726;  
RA She Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,  
RA Aweye M.J., Chan-Weiher C.C.-Y., Clausen I.G., Curtis B.A.,  
RA De Moers A., Erasus G., Fletcher C., Gordon P.M.K.,  
RA Heikamp-de Jong I., Jeffries A.C., Kozeza C.J., Medina N., Peng X.,  
RA Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,  
RA Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,  
RA Garrett R.A., Ragan M.A., Sensen C.W., Van der Oost J.;  
RT "The complete genome of the crenarchaeon Sulfolobus solfataricus P2.";  
RL Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).  
DR EMBL: AE006677; AAK40778.1;  
DR InterPro: IPR002305; tRNA-synt\_1b.  
DR InterPro: IPR002306; tRNA-synt\_1tp.  
DR Pfam: PF00579; tRNA-synt\_1b; 1.  
DR PRINTS: PR01039; TRNASYNTHTRP.  
KW Aminoacyl-tRNA synthetase; Ligase; Complete proteome.  
SQ SEQUENCE 386 AA; 45448 MW; 3F4DA8C9B8D219DF CRC64;

Query Match 42.9%; Score 907; DB 17; Length 386;

Best Local Similarity 48.6%; Pred. No. 2.8e-71;  
Matches 191; Conservative 68; Mismatches 110; Indels 24; Gaps 10;

Qy 8 TPAEDPDPWTQVTSKAG-IDYDKLIVRGSSKIDKELINRIERATGQRPHNLRGIGFS 66  
Db 6 TMDPEFTVPMVEV-----KGKVDYDKLIVQFGTKITSELEKIKSIINDELHVMRLRDVFFS 82  
Qy 67 FSHRDMQVLDAYENKRPFLYITGRGSSSEAMHGHILPFTKWLQDVENVPLVIQMT 126  
Db 60 FSHRDLVLVDYENKRPFLYITGRAPSL-GMHIGHILPFTKWLQDVENVPLVIETDE 118  
Qy 127 DDEKYLWK-DLTLDAYGDAVENAKDIIACGFDINKTFISDLDMGSSGFEKYNVVIQ 185  
Db 119 DDEKYLWK-DLTLDAYGDAVENAKDIIACGFDINKTFISDLDMGSSGFEKYNVVIQ 185  
Qy 186 KHTFENQYKGIFFGTFSDICIGKISFPALOGAOTKMSADPNSSIFLTPAKOIKTIV 245  
Db 176 KHTFENQYKGIFFGTFSDICIGKISFPALOGAOTKMSADPNSSIFLTPAKOIKTIV 245  
Qy 246 YRMTDVAAPRIGYPPKALHSTFPALOGAOTKMSADPNSSIFLTPAKOIKTIV 305  
Db 226 YRMTDVAAPRIGYPPKALHSTFPALOGAOTKMSADPNSSIFLTPAKOIKTIV 305  
Qy 306 AFSGRDITEEHRQFGNCDVVSFMYLTFLEDD-KLEQIRKDYSGAMLTGELKALIV 364  
Db 286 AFSGRDITEEHRQFGNCDVVSFMYLTFLEDD-KLEQIRKDYSGAMLTGELKALIV 364  
Qy 365 IEVQPLIAEHQARRKEVTEIVKEMPRKLS 397  
Db 346 IEVQPLIAEHQARRKEVTEIVKEMPRKLS 397

## RESULT 12

Q901F5 PRELIMINARY; PRT; 490 AA.  
AC Q901F5;  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DE 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE TRYPHOPHANYL-TRNA SYNTHETASE.  
GN L1063.04.  
OS Leishmania major.

QY	64	RGIFSHSHDMMQVLDAYENKKRPYLITGRGSSSEAMNHGHLIPFLFTKMLDQVNVPLVI	123
QY	64	RGIFSHSHDMMQVLDAYENKKRPYLITGRGSSSEAMNHGHLIPFLFTKMLDQVNVPLVI	123
Db	67	RSVFPAHEDNFENIDLAYEGRDKFYLYGRGSSSEALHGHILPMTFTYLYLOEAERKVP_LVI	126
QY	124	QMTDEKXLMKDLTDQAYGVDAVENAKOIIICGPFINTFTFSFDLDYDKMGSSGFYKNVVK	183
Db	127	QLTDEKXIMKNLSVEESQRLARENAKIIICGPDVYTFTEFSDPDYGV--GATFYKNVVK	184
QY	184	IQKHTEFNQVKGIFGFTDSDICIGKISFPAIAQAAFSNSFPQIFRDRTDIOCLIPCAIDQ	243
Db	185	VKGCVTLKAMGIFGSGEDPELAKISFPVQAVPSFPSPFHLPEFGKDNLCILPCAIQDQ	244
QY	244	DPEFMETDVAAPRIGYKPPALLHSFFPALLOGAQTMSASPNSSIFLTDPKAKOIKTVN	303
Db	245	DPEFMETDVAAPRIGYKSPALIESFFPALLOGENKMSASPNSSAIYVTDASAKOIKTVN	304
QY	304	KHAFSGRGTLEENHOFQNGCNDVPSFMUYLFPELDDDKLQIKRDYSGAMLTGELKA	363
Db	305	KYAFSGGDSLEKHRELKANLEVDIPVYLYLSFPLEDSELEHIKKEYGEGMLTGEVKKR	364
QY	364	LIEVLQPLIAEHQARRKEVDIEIKFMTPKLSFDQ	401
Db	365	LTEVLTEIVERHRRARAATVDEMVDAPVAAPRLPSMFE	402
RESULT 9			
Q09UIR2			
ID	Q09UIR2	PRELIMINARY;	PRT; 324 AA.
AC	Q09UIR2:		
DT	01-MAY-2000 (TReMBLrel. 13, Created)		
DT	01-MAY-2000 (TReMBLrel. 13, Last annotation update)		
DT	01-JUN-2001 (TReMBLrel. 17, last annotation update)		
DE	Y80D3A.1 PROTEIN.		
GN	Y80D3A.1.		
OS	Caenorhabditis elegans.		
OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdilitidae; Rhabditoidea;		
OC	Rhabdilitidae; Pelodierinae; Caenorhabditis.		
OX	NCBI_TaxID=6239;		
RA	SEQUENCE FROM N.A.		
RA	Harris B.R.:		
RL	Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.		
RN	(2)		
RP	SEQUENCE FROM N.A.		
RP	MEDLINE=99069613; PubMed=9851916;		
RA	none;		
RT	"Genome sequence of the nematode C. elegans: A platform for		
RT	investigating biology";		
RL	Science 282:2012-2018(1998).		
DR	EMBL; AL132853; CAB60439.1; -		
DR	InterPro; IPR002305; tRNA-synt_1b.		
DR	InterPro; IPR002306; tRNA-synt_trp.		
DR	Pfam; PF00579; tRNA-synt_1b; 1.		
DR	PRINTS; PR01039; TRNASYNTTRP.		
SO	SEQUENCE 324 AA; 36289 MW; 6687B6E420EC12 CRC64;		
Query Match			
	Best Local Similarity	48.5%; Score 1027; DB 5; Length 324;	
	Matches 194; Conservative 39; Mismatches 64; Indels 2; Gaps 1;		
QY	4	GPDATEAEDEYDPTVQTSKAKIDYDKLIVRGSSSIDKELINIRIRATGQRPHEFLR	63
Db	23	GGGVQDEDEDRVTPPEVYTTTKATGIDYDKLIVKFGSCRLDEIILAFERYVGNKASPMRL	82
QY	64	RGIFSHSHDMMQVLDAYENKKRPYLITGRGSSSEAMNHGHLIPFLFTKMLDQVNVPLVI	123
Db	83	RGMFPAHNDLALIDRKQGRPYLYTGRGSSSLHGLHVPFLFTKMLDQVNVPLVI	142
QY	124	QMTDEKXLMKDLTDQAYGVDAVENAKOIIICGPFINTFTFSFDLDYDKMGSSGFYKNVVK	183
Db	143	QMTDEKXLMKDKAKDEKKAREMKKIIISGDPPTTFTEFNNDY--KCPPEFENLVK	200
QY	184	IQKHTEFNQVKGIFGFTDSDICIGKISFPAIAQAAFSNSFPQIFRDRTDIOCLIPCAIDQ	243



DR InterPro: IPR002305; tRNA-synt\_1b.  
DR InterPro: IPR001412; tRNA-synt\_1.  
DR InterPro: IPR002306; tRNA-synt\_trp.  
DR Pfam: PF00579; tRNA-synt\_1b; 1.  
DR PRINTS: PR01039; tRNA-syntHTRP.  
DR PROSITE: PS00178; AA-TRNA\_LIGASE\_I; 1.  
SQ SEQUENCE 329 AA; 37613 MW; 701E702DC244C4A2 CRC64;

Query Match 77.0%; Score 1629; DB 11; Length 329;  
Best Local Similarity 92.1%; Pred. No. 9.6e-135;  
Matches 303; Conservative 17; Mismatches 9; Indels 0; Gaps 0;

QY 73 MNQLDAYENKKPFYLTGGRSPSEAMHVGHLIPFIFTKWLDVFNPLVLTQMTDDEKYL 132  
DB 1 MNQLDAYENKKPFYLTGGRSPSEAMHVGHLIPFIFTKWLDVFNPLVLTQMTDDEKYL 60  
QY 133 WKDLTLDQADAVENAKDITACGPDINKTFIFSDLDYMGSSGFYKNVYIKQHVFNQ 192  
DB 61 WKDLTLDQADAVENAKDITACGPDINKTFIFSDLEYMGSSPGFYRNVYIKQHVFNQ 120  
QY 193 VKGIFGFTDSCIGKISFPALQAPSFNSPQIFRORTDIQCLIPCAIDDPFRMTRD 252  
DB 121 VKGIFGFTDSCIGKISFPALQAPSFNSPQIFRORTDIQCLIPCAIDDPFRMTRD 180  
QY 253 VAPRIGYPRKALLHSTFFPALQAGQTKMSADPNSSIFLTPAKOIKTKVKNHAFSGRD 312  
DB 181 VAPRIGYPRKALLHSTFFPALQAGQTKMSADPNSSIFLTPAKOIKTKVKNHAFSGRD 240  
QY 313 TIEHROFGNCVDVFSFMYLTFFLEDDEKLEQIRKDYTSGAMLTGELKALIEVLQPLI 372  
DB 241 TIEHROFGNCVDVFSFMYLTFFLEDDEKLEQIRKDYTSGAMLTGELKALIEVLQPLI 300  
QY 373 AEHQARREYTDDELVEKFMTPRKLSFDEQ 401  
DB 301 AEHQARREYTDDELVEKFMTPRKLSFDEQ 329

## RESULT 4

070184 PRELIMINARY; PRT; 305 AA.

AC 070184;  
DT 01-AUG-1998 (Tremblrel. 07, Created)  
DT 01-AUG-1998 (Tremblrel. 07, Last sequence update)  
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)  
DE TRYPOPHAN-tRNA SYNTHETASE (FRAGMENT).  
OS Cavia porcellus (Guinea pig).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Hystriocognathi; Cavidae; Cavia.  
OX NCBI\_TaxID=10141;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=HARTLEY; TISSUE=SPLEEN;  
RA Yang D., Goto R., Watanabe N., Kobayashi Y.;  
RT Identification and Cloning of Genes Whose Expressions are Elevated  
RT during DNCB-induced Guinea Pig Skin Delayed-type Hypersensitivity  
RT Reaction.";  
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.  
EMBL: AB012322; BAA25288.1; -  
DR InterPro: IPR002305; tRNA-synt\_1b.  
DR InterPro: IPR002306; tRNA-synt\_trp.  
DR Pfam: PF00579; tRNA-synt\_1b; 1.  
DR PRINTS: PR01039; tRNA-syntHTRP.  
KW Aminoacyl-tRNA synthetase.  
FT NON\_TER 1  
SQ SEQUENCE 305 AA; 34838 MW; 0937164333780EB7 CRC64;

Query Match 72.6%; Score 1537; DB 11; Length 305;

Best Local Similarity 94.4%; Pred. No. 1e-126;  
Matches 288; Conservative 11; Mismatches 6; Indels 0; Gaps 0;

QY 97 EAMHVGHLIPFIFTKWLDVFNPLVLTQMTDDEKYLWKDLTLDQADAVENAKDITACG 156  
DB 1 EAMHVGHLIPFIFTKWLDVFNPLVLTQMTDDEKYLWKDLTLDQADAVENAKDITACG 60

QY 157 FDIKNTFISDLDMGSSGFYKNVYIKQHVFNNOYKIGFETDSDICGISFPALQAA 216  
DB 61 FDIKNTFISDLDMGSSGFYKNVYIKQHVFNNOYKIGFETDSDICGISFPALQAA 120  
QY 217 PSFNSPQIFRDRDTDIQCLIPCAIDDPFRMTDVAIPRIGYPRKALLHSTFFPALQAA 276  
DB 121 PSFNSPQIFRDRDTDIQCLIPCAIDDPFRMTDVAIPRIGYPRKALLHSTFFPALQAA 180  
QY 277 QTKMSADPNSSIFLTPAKOIKTKVKNHAFSGGRDITIEHROFGNCVDVFSFMYLTFF 336  
DB 181 QTKMSADPNSSIFLTPAKOIKTKVKNHAFSGGRDITIEHROFGNCVDVFSFMYLTFF 240  
QY 337 LEDDDKLEQIRKDYTSGAMLTGELKALIEVLQPLIEHQARREYTDDELVEKFMTPRK 396  
DB 241 LEDDDKLEQIRKDYTSGAMLTGELKALIEVLQPLIEHQARREYTDDELVEKFMTPRK 300  
QY 397 SFDEQ 401  
DB 301 SFDEQ 305

## RESULT 5

0904Y0 PRELIMINARY; PRT; 420 AA.

ID 0904Y0;  
AC 0904Y0;  
DT 01-MAY-2000 (Tremblrel. 13, Created)  
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)  
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)  
DE TRYPOPHANYL-tRNA SYNTHETASE (FRAGMENT).  
GN AATS-TRP OR CG9735.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99250164; PubMed=10233165;  
RA Seshiah P., Andrew D.J.;  
RT "WRS-85D: A tryptophanyl-tRNA synthetase expressed to high levels in  
RT the developing Drosophila salivary gland.";  
RL MOL. Biol. Cell 10:1595-1608(1999).  
DR EMBL: AF125157; AAF20167.1; -  
DR FlyBase: FBgn0010803; Aats-trp.  
DR InterPro: IPR002305; tRNA-synt\_1b.  
DR InterPro: IPR001412; tRNA-synt\_1.  
DR InterPro: IPR002306; tRNA-synt\_trp.  
DR Pfam: PF00579; tRNA-synt\_1b; 1.  
DR PRINTS: PR01039; tRNA-syntHTRP.  
DR PROSITE: PS00178; AA-TRNA\_LIGASE\_I; 1.  
KW Aminoacyl-tRNA synthetase.  
FT NON\_TER 1  
SQ SEQUENCE 420 AA; 46975 MW; 4AF7088AF426AB6 CRC64;

Query Match 65.0%; Score 1374.5; DB 5; Length 420;  
Best Local Similarity 64.8%; Pred. No. 2.9e-112;  
Matches 254; Conservative 61; Mismatches 76; Indels 1; Gaps 1;

QY 7 ATEAEDFVDPWTVOTSSAGIDYDKLIVFGSSKIDKELINRERATGQRPHHFLRGI 66  
DB 28 ATEAEDFVDPWTVOTSSAGIDYDKLIVFGSSKIDKELINRERATGQRPHHFLRGI 87  
QY 67 FFSHRDMNQVLDAENKKPFYLTGGRSPSEAMHVGHLIPFIFTKWLDVFNPLVLTQMT 126  
DB 88 FFSHRDLHTLTLTLEQGRKPFYLTGGRSPSGSLVHGLVPIFMTKWLETFEDVPLVLTQMT 147  
QY 127 DDEKYLWKDLTLDQADAVENAKDITACGPDINKTFIFSDLDYMGSSGFYKNVYIKQ 166  
DB 148 DDEKYLWKDLTLDQADAVENAKDITACGPDINKTFIFSDLDYMGSSGFYKNVYIKQ 207  
QY 187 HATFNQKIGFETDSDICGISFPALQAPSFNSPQIFRORTDIQCLIPCAIDDPY 246  
DB 187 HATFNQKIGFETDSDICGISFPALQAPSFNSPQIFRORTDIQCLIPCAIDDPY 246





GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 24, 2002, 12:49:45 ; Search time 21.7502 Seconds  
(without alignments)  
3189.442 Million cell updates/sec

Title: US-09-813-718-10\_COPY\_71\_471  
Perfect score: 2116  
Sequence: 1 SNHGPDATAEEDFVDPMTV.....VTDEIVKEFWTPKLSFDQ 401

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Listing first 45 summaries

SPPREMBL\_19:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query	Length	DB	ID	Description
1	1957	92.5	475	11	Q9DC65	Q9DC65 mus muscula
2	1957	92.5	481	11	Q9J958	Q9J958 mus muscula
3	1629	77.0	329	11	Q9DBR9	Q9DBR9 mus musculu
4	1537	72.6	305	11	Q70184	Q70184 cavia porce
5	1374.5	65.0	420	5	Q9U4Y0	Q9U4Y0 drosophila
6	1374.5	65.0	430	5	Q9U4Y1	Q9U4Y1 drosophila
7	1370.5	64.8	430	5	Q9VHG2	Q9VHG2 drosophila
8	1301	61.5	402	10	Q9SR15	Q9SR15 arabidopsis
9	1027	48.5	324	5	Q9UIR2	Q9UIR2 caenorhabdi
10	910.5	43.0	406	17	Q976M1	Q976M1 sulfolobus
11	907	42.9	386	17	Q972X0	Q972X0 sulfolobus
12	795.5	37.6	490	5	Q9UIF5	Q9UIF5 leishmania
13	682	32.2	136	6	Q9S295	Q9S295 sus scrofa
14	626.5	29.6	301	17	Q59584	Q59584 pyrococcus
15	534.5	25.3	380	17	Q9NH66	Q9NH66 halobacteri
16	310	14.7	111	5	Q95Y18	Q95Y18 encephalit

17	299.5	14.2	426	17	Q978Y8	Q978Y8 thermoplas
18	296.5	14.0	136	6	Q9TS88	Q9TS88 bos taurus
19	294	13.9	157	5	Q9U533	Q9U533 trypanosoma
20	286	13.5	109	4	Q9UD15	Q9UD15 homo sapien
21	269.5	12.7	513	17	Q9HN83	Q9HN83 halobacteri
22	268.5	12.7	426	17	Q9HIW5	Q9HIW5 thermoplas
23	188	8.9	364	17	Q9YA64	Q9YA64 aeropyrum p
24	185.5	8.8	341	16	Q97NA2	Q97NA2 streptococ
25	182.5	8.6	340	16	Q99XH4	Q99XH4 streptococ
26	181.5	8.6	331	17	Q99XZ1	Q99XZ1 thermoplas
27	174.5	8.2	341	16	Q9C1D1	Q9C1D1 lactococcus
28	167	7.9	351	16	Q9RVD6	Q9RVD6 delnococcus
29	164.5	7.8	895	10	Q9SGN2	Q9SGN2 arabidopsis
30	163	7.7	460	10	P93018	P93018 arabidopsis
31	158.5	7.5	375	17	Q9V0Z7	Q9V0Z7 pyrococcus
32	154.5	7.3	102	1	007119	007119 halobacteri
33	152.5	7.2	682	5	Q9N9B8	Q9N9B8 leishmania
34	149.5	7.1	327	17	Q9HN62	Q9HN62 halobacteri
35	149.5	7.1	408	10	P93363	P93363 nicotiana t
36	146	6.9	375	17	Q58739	Q58739 pyrococcus
37	145	6.9	332	17	Q9HK73	Q9HK73 thermoplas
38	142.5	6.7	528	11	Q91WQ3	Q91WQ3 mus musculu
39	139	6.6	294	17	Q96IV3	Q96IV3 sulfolobus
40	135	6.4	365	16	Q9KNV7	Q9KNV7 vibrio chol
41	134.5	6.4	525	5	Q9VW60	Q9VW60 drosophila
42	130.5	6.2	419	16	Q92BB1	Q92BB1 listeria in
43	129	6.1	339	2	Q9KZ47	Q9KZ47 streptococ
44	124	5.9	344	16	Q9AC05	Q9AC05 caulobacter
45	123	5.8	347	10	082313	082313 arabidopsis

## ALIGNMENTS

## RESULT 1

ID Q9DC65 PRELIMINARY; PRT: 475 AA.  
AC Q9DC65;  
DT 01-JUN-2001 (TREMREL. 17, Created)  
DT 01-JUN-2001 (TREMREL. 17, Last sequence update)  
DE ADULT MALE LUNG CDNA, RIKEN FULL-LENGTH ENRICHED LIBRARY,  
CLONE:1200002C07, FULL INSERT SEQUENCE.  
GN WARS.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=LUNG;  
RX MEDLINE=21085660; Pubmed=11217851;  
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
RA Arikawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
RA Alzawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K.,  
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
RA Kuehl P., Lewis S., Matsu Y., Nikaido I., Pesole G., Quackenbush J.,  
RA Schirral L.M., Staabli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
RA Blake J., Boffelli D., Bojunga N., Carlini P., de Bonaldo M.F.,  
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya K., Lee N.H.,  
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
RA Nordore P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,  
RA Wyszewski-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohzuki S.,  
RA Hayashizaki Y.,  
RT "Functional annotation of a full-length mouse cDNA collection.";  
RL Nature 409:685-690(2001).  
DR EMBL; AK004541; BAB23357.1; -.



```

QY 219 FSNFPOIFRDRDIOCLIPCAIDODPYFRMTRVAPRIGYKPPALHSTFFPAL-OGAQ 277
      : | | : | | : | | : | | : | | : | |
Db 176 -----LDEGFLDVDCQFG-GVDQRKIFVLAENLPSLGKKRAHLMNPMVPGLAQCG- 226
      : | | : | | : | | : | | : | | : | |
QY 278 TKMSASDPNSSIFLDTAKQIKTVNKKHAFSGGRDTIEH-----ROFGG 322
      : | | : | | : | | : | | : | | : | |
Db 227 -KMSASDPNSKIDLEEKQYKKKINSAFCSPG--NVEENGILSFVQYVIAPIQELKFGT 283
      : | | : | | : | | : | | : | | : | |
QY 323 NCDVDVSEFMYLTFLEDDDK-----LEQIRKDYTSGAMLTGELK---KALIEYLO 369
      : | | : | | : | | : | | : | | : | |
Db 284 N-----HFEFIDREKFGGPITYKSFEEEMKLAFAEKKLSPPDLKIGVADAINELE 335
      : | | : | | : | | : | | : | | : | |
QY 370 PLIAEHOARRKEVTDEIVKEF--MTPRK 395
      : | | : | | : | | : | | : | | : | |
Db 336 P-IRQEFANNKEFQEAASEKGIYPATPOK 362

```

Search completed: October 24, 2002, 12:51:44  
 Job time : 8.35218 secs







```

Db 340 AYERHVEFLKDLKEKREQAELAVK 364
RESULT 9
SYM_METHH STANDARD; PRT; 364 AA.
ID SYM_METHH STANDARD; PRT; 364 AA.
AC 026352;
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA Ligase)
DE (TrpRS).
OS TRPS OR MTH251.
OS Methanobacterium thermoautotrophicum.
OC Archaea; Euryarchaeota; Methanobacteriales; Methanobacteriaceae;
OC Methanothermobacter.
OX NCBI_TaxID=145262;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-DELTA H:
RA MEDLINE=98037514; PubMed=9371463.
RA Smith D.R., Doucette-Stamm L.A., Delonghery C., Lee H.-M., Dubois J.,
RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
RA Harrison D., Hoang L., Keagle P., Lumm W., Pochler B., Qiu D.,
RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
RA Jiwani N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,
RA McDougall S., Shimer G., Goyal A., Pietrowski S., Church G.M.,
RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;
RT "Complete genome sequence of Methanobacterium thermoautotrophicum
RT delta: functional analysis and comparative genomics.";
RL J. Bacteriol. 179:7135-7155(1997).
CC -1- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +
CC diphosphate + L-tryptophanyl-tRNA(Trp).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AE000812; AAB84757.1; -.
DR InterPro: IPR002305; tRNA-synt_1b.
DR InterPro: IPR001412; tRNA-synt_1.
DR InterPro: IPR002306; tRNA-synt_trp.
DR Pfam: PF00579; tRNA-synt_1b; 1.
DR PRINTS: PR01039; TRNASYNTTRP.
DR PROSITE: PS00178; AA_TRNA_LIGASE_I; FALSE_NEG.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW Complete proteome;
FT SITE 70 "HIGH" REGION.
FT SITE 251 "KMSK" REGION.
FT SITE 255 "KMSK" REGION.
SQ SEQUENCE 364 AA; 41301 MW; C2F348903338FE61D CRC64;
Query Match 18.8%; Score 397.5; DB 1; Length 364;
Best Local Similarity 27.9%; Pred. No. 1.6e-25;
Matches 107; Conservative 76; Mismatches 153; Indels 47; Gaps 14;
QY 15 VDPPTVGTSSAKGIDYKLYRGSSSKIDKELIRERATGQRPHNHLRGIFPSHRMN 74
Db 2 IDPP-----GSAK-LEYODLIENFGVRPP-SEVLDEV-----PEPSWLMRKGIIFGHRUYE 50
QY 75 QVLDAVENKKPFYLYTGRGSSSEAMHVGHLIPFTFKMLQDVFNVPVLYTCQTDDEKYLK 134
Db 51 RIISAMKGEFAYVTGMPGSR-MHGHKMLIVQQLRW-YRMGAELFIPIADMEATYSAR 108
QY 135 DLTLDQAVGAIVEN-ANDTIACGFDINK-----TTFPSDLIDWGMSSGFFYKKNVYK 183
Db 109 GVDEDSRRIRAIIEEYIAGYIALGLDEKDNITHVYLQSENLMAVEDLAYV----- 156

```

---

```

QY 184 IQKHVFENQVKGIFGTFDSDICGISFPALQAAFSNSFPQIFRDTIOCLIPCAIDQ 243
Db 157 LAGKVNENELRAITGFGSTSMAMHVAPIIIVSDILHRQDELGLGPR---PVIVPGEDQ 213
QY 244 DVFYFMRDVAARI---GYPKPALHSTFPFALQAGTQKMSASDPNSIFLTDTAKQIK 299
Db 214 DPHIRLRLDIAARFRDRGFIPLSSYTHRFMGGLTGG--KMSNRPKSAIFLSDTPEEAE 271
QY 300 TKVKKHAFSGGRTIIEHRQFGNCDVDVSNYLTFLFE-DDDLLEQIRKRYTSGAMLTG 358
Db 272 AKI-RNAKGTGRETLKQRELGVPDECIIYETLLYHMSGSDSLREIYESCRNGTLMCG 330
QY 359 ELKKALIEVLDPLIAEQHARRKE 381
Db 331 ECKNNYAEFIKKFPEELSVKREK 353
RESULT 10
SYM_ENCCU STANDARD; PRT; 134 AA.
ID SYM_ENCCU STANDARD; PRT; 134 AA.
AC 096771;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA Ligase)
DE (TrpRS) (Fragment).
OS Eucaryota; Eukaryota; Unikaryonidae; Eucaryotaxon.
OX NCBI_TaxID=6035;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98277683; PubMed=9615449.
RA Peyrethallade E., Broussolle V., Peyret P., Metenier G., Gouy M.,
RA Vivares C.P.;
RT "Microsporidia, amitochondrial protists, possess a 70-kDa heat shock
RT protein gene of mitochondrial evolutionary origin.";
RL Mol. Biol. Evol. 15:683-689(1998).
CC -1- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +
CC diphosphate + L-tryptophanyl-tRNA(Trp).
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AJ012470; CAA10034.1; -.
DR HSP: P00952; ITC.
DR InterPro: IPR001412; tRNA-synt_1.
DR PROSITE: PS00178; AA_TRNA_LIGASE_I; 1.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding.
FT SITE 89 "HIGH" REGION.
FT SITE 134 "HIGH" REGION.
SQ SEQUENCE 134 AA; 15744 MW; 60E2935B7E1E344F CRC64;
Query Match 18.2%; Score 386; DB 1; Length 134;
Best Local Similarity 51.5%; Pred. No. 3.7e-25;
Matches 68; Conservative 25; Mismatches 35; Indels 4; Gaps 1;
QY 12 EDEVDPTVGTSSAK-----GIDYKLYRGSSSKIDKELIRERATGQRPHNHLRGIF 67
Db 3 EORTIPMVVEVSTDEVDVAIDYKIIQFCEKFNQALDRLEKLSGKPAHYEFFRRIV 62
QY 68 FSHDMMQVLDAYENKKPFYLYTGRGSSSEAMHVGHLIPFTFKMLQDVFNVPVLYTCQTD 127
Db 63 FAHDFNLLDELIANNRPFLYLTGRGSSKTMHGHITFIPLCKYMODAFKIRLYQITLD 122
QY 128 DEKYLWKDLTLD 139

```



"Pyrococcus abyssi genome sequence: insights into archaeal chromosome structure and evolution."  
RT Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.  
CC -1- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +  
CC diphosphate + L-tryptophanyl-tRNA(Trp).  
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: AJ248288; CAB50601.1; -  
DR InterPro: IPR002305; tRNA-synt\_1b.  
DR InterPro: IPR001412; tRNA-synt\_1.  
DR InterPro: IPR002306; tRNA-synt\_trp.  
DR Pfam: PF00579; tRNA-synt\_1b; 1.  
DR PRINTS: PR01039; tRNAsyntHTRP.  
KM Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;  
KW Complete proteome.  
FT SITE 82 "HIGH" REGION.  
FT SITE 253 257 "KMSKS" REGION.  
SQ SEQUENCE 385 AA; 45100 MW; 4C29D0414976B12 CRC64;  
Query Match 37.9%; Score 803; DB 1; Length 385;  
Best Local Similarity 45.3%; Pred. No. 5.1e-59;  
Matches 173; Conservative 63; Mismatches 120; Indels 26; Gaps 9;  
QY 12 EDF-VDPMTQTSASAKIDYDKLIVRFSGSKIDKELINRIERATGPRPHFLRGIFESH 70  
DB 3 EDFVTPWMEVGV---VDYNKLEHFGTSTLPELLEKTAELTKSLPLEFRKKFFSH 58  
QY 71 RDNQVLDAVENKKPFLYLYGRGPSSAMHVGHLIPFLTKLQDVENVPLVQMTDEK 130  
DB 59 RDXKVLQDVEEGRGFLYGRGPSG--PMHGHITPEFAFKWLOEKGVNLYIQITDEK 117  
QY 131 YLMKD-LTLQAVGDAYENAKDIAGCPDINKTFISDLDVMGSSGFYKNVYIOKHVT 189  
DB 118 FLRENLTFEEDTKMAENELIDITAVGDPDKTIFQNSEF---TYIEVAIPITAKIN 173  
QY 190 FNOYKGFTEFSDICIGISPAIOAPSFNSFPQIFRDRTDIOCLIPCAIDDPYFRM 249  
DB 174 FSAKAAVFGTEGSKIGMIFPAQIAPT-----FEKR---KCLIPAIIDDPYWR 223  
QY 250 TRDVAPRIGYPPKPALHSTFPALQAGQTKMSASDPNSSIFLDTAKOIKTKVKNHAFSG 309  
DB 224 QRDFAESLGYKTAHISKEVPSLTSLGKMSASKPETATYLTDSPRDVEKKVYK 283  
QY 310 GROTIEHROFGCNDVDSFMVLTFFLEDDDKLEQIRKDY---TSGAMLTGELKALIE 366  
DB 284 GRPLTKQOREGGPEKCVVFKWLEIFFEEDK--KLKERIYACKNGELTGCGECKRYLIS 341  
QY 367 VLQPLIAEHQARKEVTEIYK 388  
DB 342 KIQEFLKEHOKRRKKAKEQIEK 363  
RESULT 8  
SYM\_METJUA  
ID SYM\_METJUA STANDARD; PRT; 370 AA.  
AC 058810;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA ligase)  
DE (TrpS).  
GN TRPS OR MJ1415.  
OS Methanococcus jannaschli.  
OC Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;

Methanococcus.  
OX NCBI\_TaxID=2190;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;  
RX MEDLINE=96337999; PubMed=8688087;  
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,  
RA Sutton G.G., Blake O., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,  
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,  
RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhmann J.L., Nguyen D.,  
RA Usterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,  
RA Cotton M.G., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,  
RA Klenk H.-P., Fraser C.M., Smith H.O., Moese C.R., Venter J.C.;  
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus  
jannaschli."  
RT Science 273:1058-1073(1996).  
CC -1- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +  
CC diphosphate + L-tryptophanyl-tRNA(Trp).  
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: U67582; AAB99425.1; -  
DR TIGR: MJ1415; -  
DR InterPro: IPR002305; tRNA-synt\_1b.  
DR InterPro: IPR001412; tRNA-synt\_1.  
DR InterPro: IPR002306; tRNA-synt\_trp.  
DR Pfam: PF00579; tRNA-synt\_1b; 1.  
DR PRINTS: PR01039; tRNAsyntHTRP.  
DR POSITIVE: PS00178; AA tRNA\_LIGASE\_1; FALSE\_NEG.  
KM Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;  
KW Complete proteome.  
FT SITE 75 83 "HIGH" REGION.  
FT SITE 255 259 "KMSKS" REGION.  
SQ SEQUENCE 370 AA; 42660 MW; E6C71107CF82B59D CRC64;  
Query Match 19.4%; Score 409.5; DB 1; Length 370;  
Best Local Similarity 30.9%; Pred. No. 1.7e-26;  
Matches 119; Conservative 72; Mismatches 153; Indels 41; Gaps 16;  
QY 17 PWTVQTSASAKIDYDKLIVRFSGSKIDKELINRIERATGPRPHFLRGIFESHMDYNOV 76  
DB 8 PW-ETPAV--IDYKKTMEQGVYPIVYGLDLEE-----HHFFRNIIIGHRPERI 57  
QY 77 LDVAENKKPFLYLYGRGPSSAMHVGHLIPFLFKWQ---DVPNVPLVQMTDEKYLW 133  
DB 58 VDAIKNKEFAVVSVMGMPGSK-MHFGRKMYVDLKFQKYKTDMNIP1---ADLEAYWA 112  
QY 134 KDLTLQAVGDV-ENAKDIAGCFDINKTFISDLDVMGSSGFYKNVYI-OKHTFN 191  
DB 113 RNMSFETTKELALNEYTNTIALGLDPEKINVIQSKYQV-----KDLALISKRKNWS 167  
QY 192 QVKGIFETDSDICIGISPAIOAPSFNSFPQIFRDR--DIOCLIPCAIDDPYFRM 249  
DB 168 EMKAIYGFGEFTNIGHFAPIYVADIL--HPQDLBNLSPEKPPVYVPGVIGDDPHIRL 224  
QY 250 TRDVAPR---IGYKPALHSTFPALQAGQTKMSASDPNSSIFLDTAKOIKTKVKNHAFSG 306  
DB 225 TRDIANAKKEFKFLPSSYTHRENTGLLG--KMSSSKPEPTALFLTDEDETVKKRTFS-A 281  
QY 307 FSGGROTIEHROFGC---MCDVDSFMVLTFFLEDDDKLEQIRKDY---TSGAMLTGELKKA 363  
DB 282 KITGRETLIEHKKYGVPECVYIELPLY--HLIDDKELAIETIYQKRSGSELIGCKRKKM 339  
QY 364 LIEVQPLIAEHQARKEVTEIYK 388

KW Ligase; ATP-binding.  
 FT SITE 91 "HIGH" REGION.  
 FT SITE 279 "KMSKS" REGION.  
 SQ SEQUENCE 395 AA; 44910 MW; E656AE8B76C5FDE9 CRC64;  
 Query Match 57.28; Score 1210; DB 1; Length 395;  
 Best Local Similarity 59.68; Pred. No. 1.2e-92;  
 Matches 235; Conservative 60; Mismatches 91; Indels 8; Gaps 4;

QY 11 EEDVDPMWVOTQTS-----SAGIDYDKLIVRGSSKIDKELINLRERATGQRPHFLRGI 66  
 DB 4 EEOIVTMDVKGSTVDEEGEIGDIERLIVGTRKRTPEOLEPEREKLTKKPHLLRGA 63  
 QY 67 FFSHRDNOVDAENKRPFLYTGRRPSSSEAMHGHILPIFTKWLQDVFNPLVIQMT 126  
 DB 64 FFSHRDMDLIDREOKKPFYLTGRGPSSSMHLMHPIPFCKWLDQVQVPLVIQLT 123  
 QY 127 DDEKYLKMD-LTLDQAYGDAVENAKDIIAGCGDINKTFIESDLDMGSSGFYKVVKIQ 185  
 DB 124 DDEKFLFKOGVLEDCORFAENAKDIIAVGFDPKTFIEFNSTYVG--GAFYQNVVRIA 181  
 QY 186 KHVFNQVKGIFGFTSDSCIGKISFPALQAAFSNSPQIFRRTDIOCLIPCAIDDP 245  
 DB 182 KCIYANOSKACFGFTSDSDSGIKHFASIQAAFSFSSPHIFENAKDIPCLIPCAIDDP 241  
 QY 246 YFRMTDVARIRIGYPRKALLHSTFFPALQGAQTAKMSASDPNSIFLDTAKOIKTKYK 305  
 DB 242 YFRLTRVSGKLRKKKALLHSTRFPALQGPQSKMSASDSSALFPMIDTPKIKINKINRH 301  
 QY 306 AFSGGRDITIEHROFGNCVDVSEMYLTFELEDKLEQIRKDYTSGLMTGLKALI 365  
 DB 302 AFSGCGATIEHREKGNPNVDVAYQYLSFELDDKELKQLYNTYKAGLTSTGEMKECI 361  
 QY 366 EYIQPLAEHQARKEVTDVYKFEPM--PKLSF 398  
 DB 362 KLIQGFVSDFOARSKVDENTLDMFGDSRKLW 395

RESULT 6  
 SYNC\_YEAST  
 ID STWC\_YEAST STANDARD; PRT; 432 AA.  
 AC Q12109;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 01-NOV-1997 (Rel. 35, Last annotation update)  
 DE Tryptophanyl-tRNA synthetase, cytoplasmic (EC 6.1.1.2) (Tryptophan--  
 tRNA ligase) (TrpS).  
 GN WRS1 OR YOL097C OR HRE432.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
 OX NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96076631; Pubmed=7502582;  
 RA Vandenbol M., Durand P., Portetelle D., Hilger F.;  
 RT "Sequence analysis of a 44 kb DNA fragment of yeast chromosome XV  
 including the Tyl-H3 retrotransposon, the sufi(+)-frameshift  
 RT suppressor gene for tRNA-Gly, the yeast transfer RNA-Thr-1a and a  
 RT delta element.";  
 RL Yeast 11:1069-1075(1995).  
 RN [2]  
 RP FUNCTION  
 RX MEDLINE=97197969; Pubmed=9046085;  
 RA John T.R., Ghosh M., Johnson J.D.;  
 RT "Identification and expression of the Saccharomyces cerevisiae  
 RT cytoplasmic tryptophanyl-tRNA synthetase gene.";  
 RL Yeast 13:37-41(1997).  
 CC -1- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +  
 CC diphosphate + L-tryptophanyl-tRNA(Trp).  
 CC -1- SUBUNIT: HOMODIMER.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL: Z48149; CAAB8164.1; -;  
 DR EMBL: Z74839; CA99110.1; -;  
 DR SGD: S0005457; WRS1.  
 DR Interpro: IPR002305; tRNA-synt\_1b.  
 DR Interpro: IPR001412; tRNA-synt\_1.  
 DR Interpro: IPR002306; tRNA-synt\_trp.  
 DR Pfam: PF00579; tRNA-synt\_1b; 1.  
 DR PRINTS: PR01039; TRNASYNTHTRP.  
 DR PROSITE: PS00178; AA-TRNA\_LIGASE\_1; 1.  
 KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding.  
 FT SITE 111 120  
 FT SITE 295 "KMSKS" REGION.  
 SQ SEQUENCE 432 AA; 49350 MW; C408F169737E9736 CRC64;

Query Match 55.08; Score 1163; DB 1; Length 432;  
 Best Local Similarity 54.88; Pred. No. 1e-88;  
 Matches 217; Conservative 71; Mismatches 100; Indels 8; Gaps 3;

QY 7 ATEAEEDFVDPWVY-----QTSKAGIDYDKLIVRGSSKIDKELINLRERATGQRPHHF 61  
 DB 19 STDVKEQVYTPMDVEGGVDGGRQANDIDYDLKLFQETKVPNETTLKRFQVYGRPHHF 78  
 QY 62 LRRCIFFSHDMNOVDAYENKRPFLYTGRRPSSSEAMHGHILPIFTKWLQDVFNPL 121  
 DB 79 LRKLFSSERDFKILDLLEGGRKPFYLTGRGPSSSMHLMHPIPFCKWLDQVQVPL 138  
 QY 122 VIQMTDEKYLKMD-LTLDQAYGDAVENAKDIIAGCGDINKTFIESDLDMGSSGFYK 180  
 DB 139 VIELTDEKFLFKHKLITINDYKNFARENAKDIIAVGFDPKTFIEFSDIQYMG--GAFYET 196  
 QY 181 VVKIQKHVFNQVKGIFGFTSDSCIGKISFPALQAAFSNSPQIFRRTDIOCLIPCA 240  
 DB 197 VWRVSRQITGSTAKAVGFGNDSDCIGKFHPSIQIATAFSSFPFNVLGLDPKTPCLIPCA 256  
 QY 241 IDDDPYFRMTDVARIRIGYPRKALLHSTFFPALQGAQTAKMSASDPNSIFLDTAKOIKT 300  
 DB 257 IDDDPYFRVCRDVAADKIKSPALLHSTRFPALQGSTTKMSASDDTIAITWTDPKOIK 316  
 QY 301 KVNKHAESGGRDITIEHROFGNCVDVSEMYLTFELEDKLEQIRKDYTSGLMTGL 360  
 DB 317 KINKYARSGGQVSADLHRELGNPNVDVAYQYLSFKDDVFLKECYDKYKSGELLSGEM 376  
 QY 361 KKAILEVQLPLAEHQARKEVTDVYKFEPM--PKLSF 396  
 DB 377 KLCIETLQEFVKAFOERRAQVDETLKPMVPHKL 412

RESULT 7  
 SYN\_PYRAB  
 ID STW\_PYRAB STANDARD; PRT; 385 AA.  
 AC Q9UY11;  
 DT 01-MAR-2002 (Rel. 41, Created)  
 DT 01-MAR-2002 (Rel. 41, Last sequence update)  
 DT 01-MAR-2002 (Rel. 41, Last annotation update)  
 DE Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA ligase)  
 DE (TrpS).  
 GN TRPS OR PAB1111.  
 OS Archaea; Euryarchaeota; Thermococcales; Pyrococcus.  
 OC Archaea; Euryarchaeota; Thermococcales; Pyrococcus.  
 OX NCBI\_TaxID=29292;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RP STRAIN=ORSAY;  
 RC Helbig R.;

```

|||||
Db 315 IDDPYFMTRDVAPRIGYKPKALLHSTFFPALGQATKMSASDPNSSIFLDTAKQIKT 374
Qy 301 KVNKHAFGSGRDTIEEHROFGNCVDVDSFMYLTFLEDDDKLEDIRKDYTSGAMLTGEL 360
Db 375 KVNKHAFGSGRDTIEEHROFGNCVDVDSFMYLTFLEDDDKLEDIRKDYTSGAMLTGEL 434
Qy 361 KKALEVLOPLIAEHQARRKEVTDIEVKEFMTPRKLSDFQ 401
Db 435 KKELIDVLOPLIAEHQARRKEVTDIEVKEFMTPRKLSDFQ 475

```

## RESULT 4

```

SYM_MOUSE STANDARD: PRT: 481 AA.
ID SYM_MOUSE

```

```

AC P32921.
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA ligase)
DE (TrpRS).
GN MARS OR MRS.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95018226; PubMed=7932716;
RA Pajot B., Sarger C., Bonnet J., Garret M.;
RT "An alternative splicing modifies the C-terminal end of tryptophanyl-
RT tRNA synthetase in murine embryonic stem cells.";
RL J. Mol. Biol. 242:599-603(1994).
CC -1- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +
CC diphosphate + L-tryptophanyl-tRNA(Trp).
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1- ALTERNATIVE PRODUCTS: TWO FORMS OF THIS PROTEIN ARE PRODUCED BY
CC ALTERNATIVE SPLICING. A ISOFORM OF 475 RESIDUES, FOUND IN MOST
CC TISSUES AND A C-TERMINALLY EXTENDED FORM OF 481 RESIDUES FOUND IN
CC EMBRYONIC STEM CELLS.
CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
CC -1- SIMILARITY: CONTAINS 1 "WHEP-TRS" DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X69656; CAA49347.1; -
DR EMBL: X69657; CAA49348.1; -
DR PIR: S31461; S31461.
DR PIR: S31462; S31462.
DR MGI: MGI:104630; MARS.
DR InterPro: IPR000738; WHEP-TRS.
DR InterPro: IPR002305; tRNA-synt_1b.
DR InterPro: IPR001412; tRNA-synt_1.
DR InterPro: IPR002306; tRNA-synt_trp.
DR Pfam: PF00579; tRNA-synt_1b; 1.
DR PRINTS: PF00458; WHEP-TRS; 1.
DR PRINTS: PR01039; TRNASYNTHTRP.
DR PROSITE: PS00178; AA_TRNA_LIGASE_I; 1.
DR PROSITE: PS00762; WHEP-TRS; 1.

```

```

KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW Alternative splicing.
FT DOMAIN 23 68 WHEP-TRS.
FT SITE 168 177 "HIGH" REGION.
FT SITE 353 357 "KMSRS" REGION.
FT VAAPSLIC 476 481 MISSING (IN MAJOR ISOFORM).
SEQUENCE 481 AA: 54282 MW; B05A452C08074F52 CAC64;

```

```

Query Match 91.6%; Score 1938; DB 1; Length 481;
Best Local Similarity 90.8%; Pred. No. 9,7e-153;
Matches 363; Conservative 21; Mismatches 16; Indels 0; Gaps 0;

```

```

Qy 2 NHGDATEAEEDVADPWTQVOTSSAKGIDYDKLIVFGSSKIDKELINIERATGQRPHE 61
Db 76 NCDSDATKASEDFVDPWMTVRTSSAKGIDYDKLIVQPGSSKIDKELINIERATGQRPHE 135
Qy 62 LRRGFFSHRDMNOVLDAENKKRPVLYTSGRPSSAMHVNHLIPFTKMLQOVFNPL 121
Db 136 LRRGFFSHRDMNOVLDAENKKRPVLYTSGRPSSAMHNLGLVLPFTKMLQOVFNPL 195
Qy 122 VIOMTDEKYLMDLTLDOAGDAVENAKDIIACGFDINTKFTISDLDYMGSSGFYKNV 181
Db 196 VIOMSDDEKYLMDLTLQAGSYVENAKDIIACGFDINTKFTISDLEYMQSGFPRNV 255
Qy 182 VKIQKHTFENVKGIFFGTSDDCIGKISFPAIQAPSFNSFPQIFRDRDIOCLIPCAI 241
Db 256 VKIQKHTFENVKGIFFGTSDDCIGKISFPAVQAPSFNSFPKIFRDRDIOCLIPCAI 315
Qy 242 DDDPYFMTRDVAPRIGYKPKALLHSTFFPALGQATKMSASDPNSSIFLDTAKQIKT 301
Db 316 DDDPYFMTRDVAPRIGYKPKALLHSTFFPALGQATKMSASDPNSSIFLDTAKQIKT 375
Qy 302 VNKHAFGSGRDTIEEHROFGNCVDVDSFMYLTFLEDDDKLEDIRKDYTSGAMLTGEL 361
Db 376 VNKHAFGSGRDTIEEHROFGNCVDVDSFMYLTFLEDDDKLEDIRKDYTSGAMLTGEL 435
Qy 362 KKALEVLOPLIAEHQARRKEVTDIEVKEFMTPRKLSDFQ 401
Db 436 KTLIDVLOPLIAEHQARRKAVTEVKEFMTPRKLSDFQ 475

```

## RESULT 5

```

SYM_SCHPO STANDARD: PRT: 395 AA.
ID SYM_SCHPO

```

```

AC Q09692;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Putative tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA
DE ligase) (TrpRS).
GN SPAC2F7.13C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Gentles S., Churcher C.M., Barrell B.G., Rajandream M.A., Walsh S.V.;
RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +
CC diphosphate + L-tryptophanyl-tRNA(Trp).
CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z50142; CAA90500.1; -
DR InterPro: IPR002305; tRNA-synt_1b.
DR InterPro: IPR001412; tRNA-synt_1.
DR InterPro: IPR002306; tRNA-synt_trp.
DR Pfam: PF00579; tRNA-synt_1b; 1.
DR PRINTS: PR01039; TRNASYNTHTRP.
DR PROSITE: PS00178; AA_TRNA_LIGASE_I; 1.

```

```

KW Hypothetical protein; Aminoacyl-tRNA synthetase; Protein biosynthesis;

```

CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----  
 DR EMBL: X53918; CAA37872.1; -  
 DR EMBL: X52113; CAA36356.1; -  
 DR EMBL: M74074; AAA30799.1; -  
 DR PIR: A40279; YHEO.  
 DR PIR: S14540; S14540.  
 DR InterPro: IPR00738; WHEP-TRS.  
 DR InterPro: IPR002305; tRNA-synt\_1b.  
 DR InterPro: IPR001412; tRNA-synt\_1.  
 DR InterPro: IPR002306; tRNA-synt\_trp.  
 DR Pfam: PF00579; tRNA-synt\_1b; 1.  
 DR Pfam: PF00458; WHEP-TRS; 1.  
 DR PRINTS: PRO1039; TRNASYNTHTRP.  
 DR PROSITE: PS00178; AA\_TRNA\_LIGASE\_1; 1.  
 DR PROSITE: PS00762; WHEP-TRS; 1.  
 KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding.  
 KM DOMAIN 24 69 WHEP-TRS.  
 FT DOMAIN 1 117 DISPENSABLE TO THE CATALYTIC ACTIVITY.  
 FT SITE 169 178 "HIGH" REGION.  
 FT SITE 353 357 "KMSKS" REGION.  
 FT CONFLICT 17 17 L -> M (IN REF. 2).  
 SQ SEQUENCE 475 AA; 53729 MW; F7E531750137EB32 CRC64;

Query Match 95.98; Score 2028.5; DB 1; Length 475;  
 Best Local Similarity 95.88; Pred. No. 3.1e-160;  
 Matches 384; Conservative 8; Mismatches 8; Indels 1; Gaps 1;

QY 1 SNHPPDAEAEEDVDWPTVOTSSAKGIDYDKLVRCSSKIDKELNRIERATGQRP 60  
 DB 76 SGEGLDTEADEDEVDWPTVOTSSAKGIDYDKLVRCSSKIDKELNRIERATGQRP 135  
 QY 61 FLRRGIFFSHRDMNQVDAEENKKPFYLYTGRGSSSEAMHGHLIPFTFKWLQDVENV 120  
 DB 136 FLRRGIFFSHRDMNQVDAEENKKPFYLYTGRGSSSEAMHGHLIPFTFKWLQDVENV 195  
 QY 121 LVVQMTDEKYLKDLTLDOAYGDAVENAKDIACGFDINKTFIFSLDYGMSSGFEYKN 180  
 DB 196 LVVQMTDEKYLKDLTLDOAYGDAVENAKD-ITCGEDINTFTFSDLDYGMSSGFEYKN 254  
 QY 181 VVKIQKHVTENQVKGIFGFTSDICIGKISPPAIOAASFNSFPQIFRDRDVOCLIPCA 240  
 DB 255 VVKIQKHVTENQVKGIFGFTSDICIGKISPPAIOAASFNSFPQIFRDRDVOCLIPCA 314  
 QY 241 IDDPYFRMTADVAPRIGYPRKALHSTFPFALOGAQTKKMSASPNSIFLDTAKQIKT 300  
 DB 315 IDDPYFRMTADVAPRIGYPRKALHSTFPFALOGAQTKKMSASPNSIFLDTAKQIKT 374  
 QY 301 KVNKHAESGGHDTIEHRQEGGNCDDVSEMYLTFPLEDDDKLQIRKDYSGAMLTGEL 360  
 DB 375 KVNKHAESGGHDTIEHRQEGGNCDDVSEMYLTFPLEDDDKLQIRDYSGAMLTGEL 434  
 QY 361 KKALIEVLQPLIAEHOARRKEVTEIYKEPMTPKLSFDFQ 401  
 DB 435 KKALIEVLQPLIAEHOARRKEVTEIYKEPMTPKLSFDFQ 475

RESULT 3  
 SYM\_RABIT  
 ID SYM\_RABIT STANDARD; PRT; 475 AA.  
 AC P23612; Q28607;  
 DT 01-NOV-1991 (Rel. 20, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA ligase)  
 DE (TrpRS).  
 GN MARS.  
 OS Oryctolagus cuniculus (Rabbit).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
 OX NCBI\_TaxID=9986;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90239043; PubMed=2185472;  
 RA Lee C.C., Craigen W.J., Muzny D.M., Harlow E., Caskey C.T.;  
 RT "Cloning and expression of a mammalian peptide chain release factor  
 RT with sequence similarity to tryptophanyl-tRNA synthetases.";   
 RL Proc. Natl. Acad. Sci. U.S.A. 87:3508-3512(1990).  
 RN [2]  
 RP REVISIONS TO 169-174 AND 227-228, AND FUNCTION.  
 RX MEDLINE=94009008; PubMed=8404867;  
 RA Frolova L.Y., Dalphin M.E., Justesen J., Powell R.J., Drugeon G.,  
 RA McCaughan K.K., Kisselev L.L., Tate W.P., Haenni A.-L.;  
 RT "Mammalian polypeptide chain release factor and tryptophanyl-tRNA  
 RT synthetase are distinct proteins.";   
 RL EMBO J. 12:4013-4019(1993).  
 CC -1- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +  
 CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.  
 CC -1- SIMILARITY: CONTAINS 1 "WHEP-TRS" DOMAIN.  
 CC -1- CAUTION: WAS ORIGINALLY (REF.1) THOUGHT TO BE AN EUKARYOTIC  
 CC RELEASE FACTOR (ERF).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----  
 DR EMBL: M33460; AAA31246.1; ALT-SEO.  
 DR EMBL: U02595; AAB60257.1; -  
 DR PIR: A35904; YWRBPR.  
 DR InterPro: IPR00738; WHEP-TRS.  
 DR InterPro: IPR002305; tRNA-synt\_1b.  
 DR InterPro: IPR001412; tRNA-synt\_1.  
 DR InterPro: IPR002306; tRNA-synt\_trp.  
 DR Pfam: PF00579; tRNA-synt\_1b; 1.  
 DR Pfam: PF00458; WHEP-TRS; 1.  
 DR PRINTS: PRO1039; TRNASYNTHTRP.  
 DR PROSITE: PS00178; AA\_TRNA\_LIGASE\_1; 1.  
 DR PROSITE: PS00762; WHEP-TRS; 1.  
 KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding.  
 KM DOMAIN 23 68 WHEP-TRS.  
 FT DOMAIN 168 177 "HIGH" REGION.  
 FT SITE 353 357 "KMSKS" REGION.  
 SQ SEQUENCE 475 AA; 53799 MW; 33BC9E718FE45DC4 CRC64;

Query Match 92.78; Score 1961.5; DB 1; Length 475;  
 Best Local Similarity 91.58; Pred. No. 1.1e-154;  
 Matches 367; Conservative 20; Mismatches 13; Indels 1; Gaps 1;

QY 2 NHG-PDAEAEEDVDWPTVOTSSAKGIDYDKLVRCSSKIDKELNRIERATGQRP 60  
 DB 75 SHGPEAVDDKEDVDWPTVOTSSAKGIDYDKLVRCSSKIDKELNRIERATGQRP 134  
 QY 61 FLRRGIFFSHRDMNQVDAEENKKPFYLYTGRGSSSEAMHGHLIPFTFKWLQDVENV 120  
 DB 135 FLRRGIFFSHRDMNQVDAEENKKPFYLYTGRGSSSEAMHGHLIPFTFKWLQDVENV 194  
 QY 121 LVVQMTDEKYLKDLTLDOAYGDAVENAKDIACGFDINKTFIFSLDYGMSSGFEYKN 180  
 DB 195 LVVQMTDEKYLKDLTLDOAYGDAVENAKDIACGFDINKTFIFSLDYGMSSGFEYKN 254  
 QY 181 VVKIQKHVTENQVKGIFGFTSDICIGKISPPAIOAASFNSFPQIFRDRDVOCLIPCA 240  
 DB 255 VVKIQKHVTENQVKGIFGFTSDICIGKISPPAIOAASFNSFPQIFRDRDVOCLIPCA 314  
 QY 241 IDDPYFRMTADVAPRIGYPRKALHSTFPFALOGAQTKKMSASPNSIFLDTAKQIKT 300

RA Frolowa L.Y., Grigorjeva A.Y., Sudomocina M.A., Kisselev L.L.;  
 RT "The human gene encoding tryptophanyl-tRNA synthetase: interferon-  
 RT response elements and exon-intron organization.";   
 RL Gene 128:237-245(1993).  
 RN [7]  
 RP SEQUENCE OF 265-276; 278-296; 299-317 AND 350-365.  
 RC TISSUE-Keratinoctyes; PubMed=1286667;  
 RX MEDLINE=93162043; PubMed=1286667;  
 RA Rasmussen H.H., van Damme J., Puype M., Gesser B., Cells J.E.,  
 RA Vandeckerckhove J.;  
 RT "Microsequences of 145 proteins recorded in the two-dimensional gel  
 RT protein database of normal human epidermal keratinocytes.";   
 RL Electrophoresis 13:960-969(1992).  
 RN [8]  
 RP FUNCTION.  
 RX MEDLINE=92225128; PubMed=1373391;  
 RA Bange F.-C., Flohr T., Buwitt U., Boettger E.C.;  
 RT "An interferon-induced protein with release factor activity is a  
 RT tryptophanyl-tRNA synthetase.";   
 RL FEBS Lett. 300:162-166(1992).  
 CC -I CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +  
 CC diphosphate + L-tryptophanyl-tRNA(Trp).  
 CC -I SUBUNIT: HOMODIMER.  
 CC -I INDUCTION: BY INTERFERON GAMMA.  
 CC -I SIMILARITY: BELONGS TO CLASS-I "WHEP-TRS" DOMAIN.  
 CC -I SIMILARITY: CONTAINS 1 "WHEP-TRS" DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL colustation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@sib-sib.ch).  
 CC -----  
 DR EMBL; M77804; AAA67324.1; -;  
 DR EMBL; X59892; CAA42545.1; -;  
 DR EMBL; M61715; AAA61298.1; -;  
 DR EMBL; M62570; CAA44450.1; -;  
 DR EMBL; S82905; AAB39381.1; -;  
 DR EMBL; X67920; CAB94198.1; -;  
 DR EMBL; X67921; CAB94198.1; JOINED.  
 DR EMBL; X67922; CAB94198.1; JOINED.  
 DR EMBL; X67923; CAB94199.1; JOINED.  
 DR EMBL; X67924; CAB94199.1; JOINED.  
 DR EMBL; X67925; CAB94199.1; JOINED.  
 DR EMBL; X67926; CAB94199.1; JOINED.  
 DR EMBL; X67927; CAB94199.1; JOINED.  
 DR EMBL; X67928; CAB94199.1; JOINED.  
 DR PIR; A41706; A41706.  
 DR PIR; A41633; A41633.  
 DR PIR; JH0533; JH0533.  
 DR Aatrus/Ghent-2DPAGE; 3524; IEF.  
 DR P1C1-2DPAGE; P23381; -;  
 DR MIM; 191050; -;  
 DR InterPro: IPR000738; WHEP-TRS.  
 DR InterPro: IPR002305; tRNA-synt\_1b.  
 DR InterPro: IPR001412; tRNA-synt\_1.  
 DR InterPro: IPR002306; tRNA-synt\_1.  
 DR Pfam; PF00579; tRNA-synt\_1b; 1.  
 DR Pfam; PF00458; WHEP-TRS; 1.  
 DR PRINTS; PRO1039; TRNASYNTHTRP.  
 DR PROSITE; PS00178; AA-TRNA-LIGASE\_I; 1.  
 DR PROSITE; PS00762; WHEP-TRS; 1.  
 KM Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding.  
 FT DOMAIN 19 64 WHEP-TRS.  
 FT SITE 164 173 "HIGH" REGION.  
 FT SITE 349 353 "KMSKS" REGION.  
 FT CONFLICT 213 214 SY -> GD (IN REF. 3).  
 FT CONFLICT 424 424 A -> R (IN REF. 4).  
 SQ SEQUENCE 471 AA; 53165 MW; E96344449053A0D0 CRC64;

Query Match 99.3%; Score 2101; DB 1; Length 471;  
 Best Local Similarity 99.5%; Pred. No. 3,1e-166;  
 Matches 399; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 OY 1 SNHGPDATAEEDFPDPTVOTSSAKGIDYDKLIVRFSSKIDKELINRIEATGQRHH 60  
 DB 71 SNHGPDATAEEDFPDPTVOTSSAKGIDYDKLIVRFSSKIDKELINRIEATGQRHH 130  
 OY 61 FLRRCIFFSHRDPMNOVLAVENKRPFLYTGRCPSSEAMHVGHLIPFTFKVLQDVFNVP 120  
 DB 131 FLRRCIFFSHRDPMNOVLAVENKRPFLYTGRCPSSEAMHVGHLIPFTFKVLQDVFNVP 190  
 OY 121 LVIQMTDEKYLWKDLTLDQAYDAVENAKDIIACGPDINKTFISDDYMGSSGFYKN 180  
 DB 191 LVIQMTDEKYLWKDLTLDQAYDAVENAKDIIACGPDINKTFISDDYMGSSGFYKN 250  
 OY 181 VYKICKHTFNQVKIKIFGFTSDCIGKISFPAIQAAPSFNSFPDIFDRDIDQILICA 240  
 DB 251 VYKICKHTFNQVKIGFGTSDCIGKISFPAIQAAPSFNSFPDIFDRDIDQILICA 310  
 OY 241 IDDDYFRTNRDVARPRIGYPRKALHSTFFPALQAGOTKMSADPNSSIFLDTAKQIKT 300  
 DB 311 IDDDYFRTNRDVARPRIGYPRKALHSTFFPALQAGOTKMSADPNSSIFLDTAKQIKT 370  
 OY 301 KVNKAFFSGGRDTEIENHROFGNCDDVVSFWYLFLEDDDKLEQIRDYTGAMLTGEL 360  
 DB 371 KVNKAFFSGGRDTEIENHROFGNCDDVVSFWYLFLEDDDKLEQIRDYTGAMLTGEL 430  
 OY 361 KKALIEVLQPLAEHQARKEVTDEIVKEFMTPRKLSDFQ 401  
 DB 431 KKALIEVLQPLAEHQARKEVTDEIVKEFMTPRKLSDFQ 471  
 RESULT 2  
 SYR\_BOVIN  
 ID SYR\_BOVIN STANDARD; PRT; 475 AA.  
 AC P17248;  
 DT 01-AUG-1990 (Rel. 15, Created)  
 DT 01-MAY-1992 (Rel. 22, Last sequence update)  
 DT 01-FEB-1996 (Rel. 33, Last annotation update)  
 DE Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA ligase)  
 DE (TTPRS).  
 GN WARS.  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Bovinae; Bos.  
 OX NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RC TISSUE-Rectina;  
 RX MEDLINE=91329348; PubMed=1907847;  
 RA Garret M., Pajot B., Trezeguet V., Labouesse J., Merle M.,  
 RA Gandar J.-C., Benedetto J.-P., Sallatranque M.-L., Alterio J.,  
 RA Gueguen M., Sarger C., Labouesse B., Bonnet J.;  
 RT "A mammalian tryptophanyl-tRNA synthetase shows little homology to  
 RT prokaryotic synthetases but near identity with mammalian peptide  
 RT chain release factor";   
 RL Biochemistry 30:7809-7817(1991).  
 RN [2]  
 RP SEQUENCE OF 17-475 FROM N.A.  
 RC TISSUE-Pancreas;  
 RA Garret M., Trezeguet V., Pajot B., Gandar J.-C., Merle M.,  
 RA Gueguen M., Benedetto J.-P., Sarger C., Alterio J., La Bounessec B.,  
 RA Labouesse J., Bonnet J.;  
 RT Submitted (MAR-1990) to the EMBL/GenBank/DBJ databases.  
 CC -I CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +  
 CC diphosphate + L-tryptophanyl-tRNA(Trp).  
 CC -I SUBUNIT: HOMODIMER.  
 CC -I SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.  
 CC -I SIMILARITY: CONTAINS 1 "WHEP-TRS" DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 24, 2002, 12:49:40 : Search time 7.35218 Seconds  
(without alignments)  
2111.829 Million cell updates/sec

Title: US-09-813-718-10\_COPY\_71\_471

Perfect score: 2116  
Sequence: 1 SNHGPDATAEEDFVDPWTV.....VTDEIVKEFWTPRKLSDFQ 401

Scoring table:  
BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query %	Match Length	ID	Description
1	2101	99.3	471	1	SYW_HUMAN
2	2028.5	95.9	475	1	SYW_BOVIN
3	1961.5	92.7	475	1	SYW_RABIT
4	1938	91.6	481	1	SYW_MOUSE
5	1210	57.2	395	1	SYW_SCHPO
6	1163	55.0	432	1	SYW_YEAST
7	803	37.9	385	1	SYW_PYRAB
8	409.5	19.4	370	1	SYW_METJA
9	397.5	18.8	364	1	SYW_METJA
10	386	18.2	134	1	SYW_ENCCU
11	370.5	17.5	420	1	SYW_ARCFU
12	353	16.7	374	1	SYW_AERPE
13	192	9.1	323	1	SYW_ARCFU
14	175.5	8.3	341	1	SYW_GLOLO
15	172	8.1	394	1	SYW_YEAST
16	170.5	8.1	366	1	SYW_SULSO
17	169	8.0	395	1	SYW_AOUAE
18	166.5	7.9	346	1	SYW_CHLUR
19	163.5	7.7	353	1	SYW_BORBU
20	159.5	7.5	337	1	SYW_TREPA
21	158	7.5	346	1	SYW_CHLMU
22	152	7.2	344	1	SYW_CHLPP
23	146	6.9	326	1	SYW_HELPP
24	138	6.5	319	1	SYW_METTH
25	136.5	6.5	528	1	SYW_BOVIN
26	135.5	6.4	328	1	SYW_THEMA
27	133	6.3	328	1	SYW_BACST
28	132.5	6.3	343	1	SYW_MYCLE
29	132.5	6.3	528	1	SYW_HUMAN
30	130	6.1	326	1	SYW_HELPP
31	129	6.1	401	1	SYW_SCHPO
32	128.5	6.1	379	1	SYW_YEAST
33	128	6.0	334	1	SYW_HAEIN

34	124.5	5.9	347	1	SYW_MYCLE	P47372 mycoplasma
35	122	5.8	350	1	SYW_CAEBL	P46579 ctenorhabdi
36	118.5	5.6	343	1	SYW_GLOAB	O971d6 clostridium
37	116.5	5.5	330	1	SYW_BACSU	P21656 bacillus su
38	112.5	5.3	335	1	SYW_BUCAT	P57602 buchnera ap
39	110	5.2	306	1	SYW_METJA	O57834 methanococc
40	109.5	5.2	336	1	SYW_MCTTU	O53386 mycobacteri
41	107	5.1	330	1	SYW_RICPR	O92d76 rickettsia
42	106.5	5.0	334	1	SYW_ECOLI	P00954 escherichia
43	105.5	5.0	333	1	SYW_PASMU	P57956 pasteurella
44	105	5.0	426	1	SYW_MYCLE	O49900 mycobacteri
45	104	4.9	337	1	SYW_SYNY3	P73655 synechocyst

## ALIGNMENTS

RESULT 1	SYW_HUMAN	STANDARD:	PRT:	471 AA.
AC	P23381: P78535; Q9UDL3:			
DT	01-NOV-1991 (Rel. 20, Created)			
DT	01-DEC-1992 (Rel. 24, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan--tRNA ligase)			
DE	(TrpRS) (IFP53) (hmRS).			
GN	WARS.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	MEDLINE=92105071; PubMed=1761529;			
RA	Rudin B.Y., Anderson S.L., Xing L., Powell R.J., Tate W.P.;			
RT	"Interferon induces tryptophanyl-tRNA synthetase expression in human			
RT	fibroblasts.";			
RL	J. Biol. Chem. 266:24245-24248(1991).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RA	MEDLINE=1763065;			
RA	Flecker J., Rasmussen H.H., Justesen J.;			
RT	"Human interferon gamma potentially induces the synthesis of a 55-kDa			
RT	protein (gamma 2) highly homologous to rabbit peptide chain release			
RT	factor and bovine tryptophanyl-tRNA synthetase.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 88:11520-11524(1991).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RA	MEDLINE=92112058; PubMed=1765274;			
RA	Prolova L.Y., Sudomoina M.A., Grigorjeva A.Y., Zinovjeva O.L.,			
RT	Kisselev L.L.;			
RT	"Cloning and nucleotide sequence of the structural gene encoding for			
RT	human tryptophanyl-tRNA synthetase.";			
RL	Gene 109:291-296(1991).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RA	MEDLINE=92164636; PubMed=1537332;			
RA	Buitt U., Flohr T., Boeltger E.C.;			
RT	"Molecular cloning and characterization of an interferon induced			
RT	human cDNA with sequence homology to a mammalian peptide chain			
RT	release factor.";			
RL	EMBO J. 11:489-496(1992).			
RN	[5]			
RP	SEQUENCE OF 1-13 FROM N.A.			
RA	MEDLINE=96319944; PubMed=8724762;			
RA	Sokolova I.V., Narovilianskii A.N., Amchenkova A.M., Turpaev K.T.;			
RT	"Alternative splicing of 5'-terminal exons of the human tryptophanyl-			
RT	tRNA synthetase gene.";			
RL	Mol. Biol. (Mosk) 30:319-329(1996).			
RN	[6]			
RP	SEQUENCE OF 1-141 AND 182-471 FROM N.A.			
RC	TISSUE=Spem;			
RA	MEDLINE=93292992; PubMed=7685728;			







QY 185 OKHVENOVKGFJEFTSDDCIGKISFPALOAQSPNSNSFPOIFNRDRTIOCLIPCAIDD 244  
Db 165 -RHATLOWNYG-----EPDNWGQAFYPAVQTAHLL---LPOLVHG--EHETLVPAVDOD 213  
QY 245 PYFMRTRDVABRIGYP--KPALHSHSTFPFALOGAQTOKMSASDPNNSIFLPTJAKOIKRV 302  
Db 214 PHVYRSRDVAAKARYPVKGPCGALLMQFLPSLAG-PCKMSSS-AGVISIKLTSQSPDVRREV 271  
QY 303 NKHAFSGGRDITIEHKROFGCMCDVDVSFMYLTFELEDOD-KLEQIKDYTSGAMITGELK 361  
Db 272 RTHAVTGGRAVSEEHRAAGVPAEDVPFOYLSAFEPEDAETARIERYRAGDILSGELK 331  
QY 362 KALIEVLQPLTAEHQARKKEYTDELVRKEPMRKRISFD 399  
Db 332 DLAAADRIETFLAHHORRRALGD--VTEALDAFLRTDD 367

RESULT 11  
F64476  
tryptophan--tRNA ligase (EC 6.1.1.2) - Methanococcus jannaschii  
N:Alternate names: tryptophanyl-tRNA synthetase  
C:Species: Methanococcus jannaschii  
C:Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 21-Jul-2000  
C:Accession: F64476  
R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake,  
R.; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.;  
Ison, J.D.; Sedow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.  
Science 273, 1058-1073, 1996  
A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C.  
A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii  
A:Reference number: A64300; MUID:96337999  
A:Accession: F64476  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-370 <BL>  
A:Cross-references: GB:067582; GB:L77117; NID:g1592064; PID:AB99425.1; PID:g1592065; T  
C:Genetics:  
A:Map position: FOR1375885-1376997  
A:Start codon: GTG  
C:Superfamily: mammalian tryptophan--tRNA ligase; amino acid--tRNA ligase repeat homolog  
C:Keywords: aminoacyl-tRNA synthetase; ligase; protein biosynthesis

[illegible]

RESULT 12  
E69131  
tryptophan--tRNA ligase (EC 6.1.1.2) - Methanobacterium thermoautotrophicum (strain D  
N/alternatn names: tryptophanyl-tRNA synthetase  
C:Species: Methanobacterium thermoautotrophicum  
C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 22-Oct-1999  
C:Accession: E69131  
R:Smith, D. R.; Doucette-Stamm, L. A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.  
J.; Qiu, D.; Spadafora, R.; Viscare, R.; Wang, Y.; Wierzboski, J.; Gibson, R.; Jiwanti,  
Kt, S.; Church, G.M.; Daniels, C.U.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.  
J. Bacteriol. 179, 7135-7153, 1997  
A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: fu  
A:Reference number: A69000; MUID:98037514  
A:Accession: E69131  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-364 <MTM>  
A:Cross-references: GB:AE000812; GB:AE000666; NID:g2621298; PIDN:AA84757.1; PID:g262  
A:Experimental source: strain Delta H  
C:Genetics:  
A:Gene: MTH251  
A:Start codon: TTG  
C:Superfamily: mammalian tryptophan--tRNA ligase; amino acid--tRNA ligase repeat homo  
C:Keywords: aminoacyl-tRNA synthetase; ligase; protein biosynthesis

```

Query Match      18.8% ; Score 397.5; DB 2; Length 364;
Best Local Similarity 27.9% ; Pred. No. 1e-24;
Matches 107; Conservative 76; Mismatches 153; Indels 47; Gaps 14;

OY      15 VDPWTVOTSSAKGIDYDKLIVFGSSKIDKELINRIEATGQRPHHLLRGIQFSHRDMN 74
          : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      2 IDPW-----GSAK-LEYOOLLINFCGVRF-SEVLDEV-----PEPSWMLRRRIIGFHDYE 50
          : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY      75 OYLAAYENKKFEYLTYTSGPSSSEAMHWGHLIFLFTKQLDQVFNWPLVIQMTDDEKLMK 134
          : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      51 RIISAMKKGEPAAVTGTMPSGR-MHIGHKMIVDLQR-YDRMGAEIIFIPADMEAVSAR 108
          : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY      135 DLTLDOAYGDAVEN-AKDIICAGFDINK-----TIFSLDIDMGSMSSGFGYKRWVK 183
          : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      109 GVDFEDSRRIALIEEYIAGYIALGIDLEKDNTHVYLQSENLMAVEDLAVY----- 156
          : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY      184 IQKHVTFNVOYVGIGFTDSDCIKISPAIOAAPSFSNFQIFRDRTDIOCLIPCAIDQ 243
          : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      157 LAGKAKNFELAIYIGFTGSTMAHMYAPIIIOVSDILHPOLDLELGPGR---PVIVPVGPD 213
          : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY      244 DPFYFMTEDVAPRI-----GYEKPALHLSTFPFALOGATKMSASDPPSSIFLTPTAFOIK 299
          : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      214 DPHIRLTLDIARPRDRKGFILPSSTYIRHFMGGLTG- -KMSNRKPSAIFLSTPPEAE 271
          : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY      300 TKVNKHAFSGRDTIEEHROFGMCDVDSFMYLTFLE- -DDDKLEQIRKDYTSGAMLT 358
          : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      272 AKI-RNATQSGRETLKEQRELGAVPEECIIYETLLHMGSGDSRLIEIYESCRNGTLMCG 330
          : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY      359 ELKKALILEVLPDLAENHQAARKE 381
          | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      331 ECKNNTAFFIRKFEELSVMREK 353
          | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

```

RESULT 13
T43806
tryphophan--tRNA ligase (EC 6.1.1.2) [imported] - Encephalitozoon cuniculi (fragment)
C:Species: Encephalitozoon cuniculi
C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 02-Sep-2000
C:Accession: T43806
R:Peptreallade, E.; Broussolle, V.; Peyret, P.; Metenier, G.; Gouy, M.; Vivares, C.P
Mol. Biol. Evol. 15, 683-689, 1998
A:Title: Microsporidia, amitochondrial protists, possess a 70-kDa heat shock protein
A:Reference number: 22693; MUID:98277683
A:Accession: T43806
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA

```

Db 346 IDKLNFLLEHR-RRREAKELVHFXYDGKLA 377

RESULT 8

C75020

tryptophanyl-tRNA synthetase (trps) PAB1111 - Pyrococcus abyssi (strain Orsay)

C:Species: Pyrococcus abyssi

C>Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 20-Jun-2000

C:Accession: C75020

R:anonymous, Genoscope

Submitted to the EMBL Data Library, July 1999

A:Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome, stru

A:Reference number: A75001

A:Accession: C75020

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-385 <RAW>

A:Cross-references: GB:AJ248288; GB:AL096836; NID:g5458960; PIDN:CAM50601.1; PID:g545911

A:Experimental source: strain Orsay

C:Genetics:

A:Gene: trps, PAB1111

C:Superfamily: mammalian tryptophan--tRNA ligase; amino acid--tRNA ligase repeat homolog

Query Match 37.9%; Score 803; DB 2; Length 385;

Best Local Similarity 45.3%; Pred. No. 8.4e-58;

Matches 173; Conservative 63; Mismatches 120; Indels 26; Gaps 9;

QY 12 EDF-VDVWVOTSSAKIDYDKLIVRGSSKIDKELINRERATGQRPHEFLRGIFPSH 70

Db 3 EDFVVTPEVEGV-VDYKNKLIHFSTPLTELEKTELKSELPLFFRRKFFFSH 58

QY 71 RDMQVLDVAVENKRPFLYTGSGSEAMHVGHLIPFTFWLQDVFNPLVIOMTDEK 130

Db 59 RDIDKVLQDVEEGGFELYTGSGSG-PMHIGHLIPPAIKWLQEKGVNLYITDDEK 117

QY 131 YLMKD-LTLDQAYGDAVENAKDIACGFDINKTFIFSDLDGMGSGGFYKNVAKIQKHT 189

Db 118 FLFRENLTPEETKHMVAENILDIAGVDPKTFIFQNSE-TRKYEMAIPIAKIN 173

QY 190 FNOVKGIFGTFDSCICIGSFPAIQAPSFNSFPQIFRDTDIQCLIPCAIDDDPYFRM 249

Db 174 FSAKAVFGEFGESKIGMIFPAIQAPTF-----FEKR--RCLIPALIDDDPYWRL 223

QY 250 TRDVAPIRGYKPKALHSTPEPALOGATKMSADPNSSTIFLTPAKOIKTKVKNHAFSG 309

Db 224 QRDASLGLYKTAHNSKFPVSLTSGKMSASKPETALVLTSPEDVEKKVKKFALTG 283

QY 310 GRDIEHRQFGNCDVDVSEFMYLTFLEDDDKLEQIRKDY--TSGAMLTGELKALIE 366

Db 284 GRPLTKQREKGEPEKCVVFKMLEIFFEEDDK--KLKERYACKNGELTGCCKRYLIS 341

QY 367 VLQPLIAHQARKEVDIETK 388

Db 342 KIOFLKEHQRRKKAKEQIEK 363

RESULT 9

G71206

tryptophan--tRNA ligase (EC 6.1.1.2) - Pyrococcus horikoshii

C:Species: Pyrococcus horikoshii

C>Date: 10-Aug-1998 #sequence\_revision 14-Aug-1998 #text\_change 21-Jul-2000

C:Accession: G71206

R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sekin

M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudooh, Y.; Yamazaki, J.; Koshida, N.; Oguchi

DNA Res. 5, 55-76, 1998

A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic a

A:Reference number: A71000; MUID:98344137

A:Accession: G71206

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-301 <RAW>

A:Cross-references: GB:AP000007; NID:g3236134; PIDN:BA031046.1; PID:g3258363

A:Experimental source: strain OT3

A:Note: this accession replaces an interim accession for a sequence replaced by Genba

C:Genetics:

A:Gene: PH1921

C:Superfamily: yeast tyrosine--tRNA ligase

C:Keywords: aminocacyl-tRNA synthetase; ligase; protein biosynthesis

Query Match 29.6%; Score 626.5; DB 2; Length 301;

Best Local Similarity 44.2%; Pred. No. 1.6e-43;

Matches 134; Conservative 56; Mismatches 92; Indels 21; Gaps 7;

QY 99 MHVGLIPFTKWLQDVFNPLVIOMTDEKYLMD-LTLDQAYGDAVENAKDIACGF 157

Db 1 MHIGHIIPFATKWLQEKFGVNLVLIQITDDEKFLFKEMLTFDDPKRMAYDILDIIVAGF 60

QY 158 DINKTFIFSDLDYMGSSGFYKNVAKIQKHTFNOVKGIFGTFSDICIGSFPAIQAP 217

Db 61 DPKDTFFIQNSE--TRKYEMAIPIAKKINFMAAVFEGTQSKGMIFPAIQAP 116

QY 218 SFSNSFPQIFRDTDIQCLIPCAIDDDPYFRMTSDVAPRIQYKPKALHSTFPALOGAQ 277

Db 117 TF-----FEKR--RCLIPALIDDDPYWKLQDFAESLGYTTALHNSKFPVSLTSL 166

QY 278 TRMSADPNSSTIFLTPAKOIKTKVKNHAFSGRDTIEHRQFGNCDVDVSEMYLTFEL 337

Db 167 GKMSASKPETALVLTSPEDVEKKVKKFTLTGGRPLTKEQREKGEPEKCVVFKMLEIFF 226

QY 338 EDDDKLEQIRKDY--TSGAMLTGELKALIEVLQPLAEHQARKEVTDIYVEFMPR 394

Db 227 EEDDK--KLKERYACKNGELTGCCKRYLISKIOEFLKEHQRRKK-AEKLVEKFFYTG 283

QY 395 KLS 397

Db 284 KLA 286

RESULT 10

G84373

tryptophanyl-tRNA synthetase (imported) - Halobacterium sp. NRC-1

C:Species: Halobacterium sp. NRC-1

C>Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 16-Feb-2001

C:Accession: G84373

R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Bergquist, B.; Pan, M.; Shukla, H.D.; Lasky

; Leitshauer, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Ja

Jung, K.H.; Alam, M.; Freitas, T.

Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000

A:Authors: Hou, S.; Daniels, G.J.; Dennis, P.P.; Omer, A.D.; Edhardt, H.; Lowe, T.M.;

A:Title: Genome sequence of Halobacterium species NRC-1.

A:Reference number: A84160; MUID:20504483

A:Accession: G84373

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-380 <STO>

A:Cross-references: GB:AE004437; NID:g10581646; PIDN:AAG20355.1; GSPDB:GN00138

C:Genetics:

A:Gene: trps2

C:Superfamily: mammalian tryptophan--tRNA ligase; amino acid--tRNA ligase repeat homo

Query Match 25.3%; Score 534.5; DB 2; Length 380;

Best Local Similarity 35.4%; Pred. No. 7.1e-36;

Matches 141; Conservative 76; Mismatches 140; Indels 41; Gaps 17;

QY 10 AEEDFVDPMTVOTSSAGIDYDKLIVRGSSKIDKELINRERATGQRPHEFLRGIFPS 69

Db 3 ADGNDVTPVAESD--LDYEKLARFGADELTDORARRP-----DHELVNRGLFYA 53

QY 70 HRDMQVLDVAVENKRPFLYTGSGSEAMHVGHLIPFTFWLQDVFNPLVIOMTDE 129

Db 54 GRVDVDFLTAGEQS----IYTGVPSSG-PHILGHAMFYFARRQDDEFGARVYPLSDDE 108

QY 130 KYLMKDLTLDQAYGDAVE-NAKDIACGFDINKTFIF--SDLDY-GMSSGFYKNVAKI 184

Db 109 KYWEKQTPAET--GDYLRANRLDLAVGPELIRIVVTRDADVLYPLATAFGADV--- 164

Db 124 DDEKFLKFGVSLDQCQFARENAKDIIAVGDPKKTFTLFMNSTYVG--GAFYQVWVRLA 181  
Qy 186 KHYFNVQKAGIFGFTDSDICIGKISPAIOAAPSFNSFPQIFRDRDTIOCLIPCAIDDP 245  
Db 182 KCIATNOSKACGCFDSDSICIGKIFASIOAAPSFSSSPHIFNGAKDIPCLIPCAIDDP 241  
Qy 246 YFRMRDVAAPRIGYKPKPALHSTFPFALOGAQTMMASADPNSSIFLPTAKOIKTKVNMK 305  
Db 242 YFRLLRDVSGRLKFKPKPALHSTRFPALOGPQSKMSASKDSAITMTDTPKTKKKNRH 301  
Qy 306 AFSGGRDTIEHRQFGNCDVDVSEFMYLTFPLEDDKLEQIRKDYTSGAMLTGELKAL 365  
Db 302 AFSGGATTEIHRKGNPNVDVAQYLSFLDDEKLEKQLYNTYKAGLTSGEMKGCI 361  
Qy 366 EYLOPLIAEHQARKEVYDEIVKEFMT-PRKLSF 398  
Db 362 KLLQFVSDFOAARSKVDATLDMFMDSRKLEW 395

## RESULT 6

S51901

tryptophan--tRNA ligase (EC 6.1.1.2) [similarity] - yeast (Saccharomyces cerevisiae)  
N:Alternate names: protein HRE332; protein 00792; tryptophanyl-tRNA synthetase  
C:Species: Saccharomyces cerevisiae  
C:Date: 05-May-1995 #sequence\_revision 03-Aug-1995 #text\_change 20-Jun-2000  
C:Accession: S51901; S59177; S66793  
R:Vandenbol, M.; Durand, P.; Portetelle, D.; Hilger, F.  
Submitted to the EMBL Data Library, January 1995  
A:Description: Sequence analysis of a 44kb DNA fragment of yeast chromosome XV including and a Delta.

A:Reference number: S51848

A:Accession: S51901

A:Molecule type: DNA

A:Residues: 1-432 &lt;VAM&gt;

A:Cross-references: EMBL:Z48149; NID:g663234; PIDN:CAA88164.1; PID:g663256

R:Vandenbol, M.; Durand, P.; Portetelle, D.; Hilger, F.

Yeast 11, 1069-1075, 1995

A:Title: Sequence analysis of a 44 kb DNA fragment of yeast chromosome XV including the a delta element.

A:Reference number: S59156; MUID:96076631

A:Accession: S59177

A:Molecule type: DNA

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-432 &lt;VAM&gt;

A:Cross-references: EMBL:Z48149; NID:g663234; PIDN:CAA88164.1; PID:g663256  
R:Durand, P.; Hilger, F.; Portetelle, D.; Vandenbol, M.  
Submitted to the Protein Sequence Database, July 1996

A:Reference number: S66791

A:Accession: S66793

A:Molecule type: DNA

A:Residues: 1-432 &lt;DNK&gt;

A:Cross-references: EMBL:Z74839; NID:g1419947; PIDN:CAA9110.1; PID:g1419948; GSPDB:GN0C

A:Experimental source: strain S28C

C:Genetics:

A:Gene: WRS1; MIPS:YOL097C

A:Map position: 15L

C:Superfamily: mammalian tryptophan--tRNA ligase; amino acid--tRNA ligase repeat homolog  
C:Keywords: ligase

F:117-120/Region: ATP-binding motif (HXGH)

Query Match 55.0%; Score 1163; DB 2; Length 432;  
Best Local Similarity 54.8%; Pred. No. 3.8e-87;  
Matches 217; Conservative 71; Mismatches 100; Indels 8; Gaps 3;

Qy 7 ATEAEEDVDPWTV-----QTSSAKGIDYDLIVRGSSKIDKELINRIERATGQRPNHF 61  
Db 19 STDVEQVQVTPWDVGGVDEQGRQAINIDYDLIKQFGTPVNEETLKRKQVGTGREPHNF 78  
Qy 62 LARGGFSSHRDNNQVLDAYENKKPPLYLTVGRGSSSAMVGLIPFIFPKMLQDVFNVP 121  
Db 79 LKKGLEFSERDFTKILDIYEQGRPFLLYTGKGPSSDSMHLGHIPIVFTKWLQEVFVDP 138

Qy 122 VIOMTDEKYLWK-DLTLDOAYGDAVENAKDIIACGFDINKTFIFSDLDYMGSSGFYKN 180  
Db 139 VIELTDEKFLFKHKLITINDVKNFARENAKDIIAVGDPKKTFTLFISDQYMG--GAFET 196  
Qy 181 VKIQRHVTENQVKGIFGFTDSDICIGKISPAIOAAPSFNSFPQIFRDRDTIOCLIPCA 240  
Db 197 VYRVSRLQTGSAKAVGFNDSDICIGKIFHAFSIOATATAPSSFPVGLPDKTPLLIPCA 256  
Qy 241 IDDDYFRMRDVAAPRIGYKPKPALHSTFPFALOGAQTMMASADPNSSIFLPTAKOIKT 300  
Db 257 IDDDYFRMRDVAAPRIGYKPKPALHSTFPFALOGSTTKMSASDPTFTAFMDTPKQIOK 316  
Qy 301 KYNKHAFSGGRDTIEHRQFGNCDVDVSEFMYLTFPLEDDKLEQIRKDYTSGAMLTGEL 360  
Db 317 KINKAFAFSGGQVSAHLHELGNPNVDVAQYLSFKRDDVDVFLKCYKYSNGELLSGM 376  
Qy 361 KKALEVLQPLIAEHQARKEVYDEIVKEFMTPRKL 396  
Db 377 KKLCTETLQEFVKAFQERRAQVDEETLDFKFWPHKL 412

## RESULT 7

C90190

tryptophanyl-tRNA synthetase (trps) [imported] - Sulfolobus solfataricus  
C:Species: Sulfolobus solfataricus  
C:Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 15-Jun-2001  
C:Accession: C90190  
R:She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Anayez, M.J.; Ch  
Jong, T.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder  
arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.  
Submitted to GenBank, April 2001  
A:Description: Sulfolobus solfataricus complete genome.

A:Reference number: A99139

A:Accession: C90190

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-386 &lt;KUR&gt;

A:Cross-references: GB:AB006641; NID:g13813608; PIDN:AAK40778.1; GSPDB:GN0155

C:Genetics:

C:Superfamily: mammalian tryptophan--tRNA ligase; amino acid--tRNA ligase repeat homo

Query Match 42.9%; Score 907; DB 2; Length 386;  
Best Local Similarity 48.6%; Pred. No. 2.7e-66;  
Matches 191; Conservative 68; Mismatches 110; Indels 24; Gaps 10;

Qy 8 TEAEEDVDPWTVQTSAGK-IDYDKLIVRGSSKIDKELINRIERATGQRPNHF 66  
Db 6 TWPDEFTVTPWEV-----GKQVDYDKLIVQFQTKITELKORIKNLAGDL-HVMLRRNV 59  
Qy 67 FFSHRDNNQVLDAYENKKPPLYLTVGRGSSSAMVGLIPFIFPKMLQDVFNVP 126  
Db 60 FFSHDDLDVLVDYKRSKGFLLYGRAPSL-GMHIGHLIPFTWLDKPKFANLYET 118  
Qy 127 DDEKYLWK-DLTLDOAYGDAVENAKDIIACGFDINKTFIFSDLDYMGSSGFYKNVYK 185  
Db 119 DDEKYMARNDEFLDQTRSAAYDNIIIDIAVGNPNKTFIFQOTEVI---RNMYPTIVVIA 175  
Qy 186 KHYTNVQKAGIFGFTDSDICIGKISPAIOAAPSFNSFPQIFRDRDTIOCLIPCAIDDP 245  
Db 176 KKLTFSEVRATFGDLASSNIGLIFYPALDIAPT-----MEKK---RCLIPACIDDP 225  
Qy 246 YFRMRDVAAPRIGYKPKPALHSTFPFALOGAQTMMASADPNSSIFLPTAKOIKTKVNMK 305  
Db 226 YKRLORDIAESIGYKKAQIHKFLPPLTGPBGKSSSPETAIYLVDPKTVKRNKY 285  
Qy 306 AFSGGRDTIEHRQFGNCDVDVSEFMYLTFPLEDD-KLEQIRKDYTSGAMLTGELKAL 364  
Db 286 AFSGGQPTIELHRRKYGNEIDVPFOWLYYFEEEDNRKIEIEERYSGKMLTGELKQL 345  
Qy 365 IEVLOPLIAEHQARKEVYDEIVKEFMTPRKLS 397



A:Residues: 1-212, 'GD', 215-471 <FRO2>  
A:Cross-references: GB:MK1715; NID:g340367; PID:AAA61298.1; PID:g340368  
A:Experimental source: fibroblast  
C:Genetics:  
A:Gene: GDB:WARS; IPP53  
A:Cross-references: GDB:119632; OMIM:191050  
A:Map position: 14q23-14q31  
A:Introns: 33/3; 105/1; 141/2; 181/2; 242/2; 276/1; 313/3; 371/3; 418/3  
C:Superfamily: mammalian tryptophan--tRNA ligase; amino acid--tRNA ligase repeat homology  
C:Keywords: aminocyl-tRNA synthetase; ATP; ligase; protein biosynthesis  
C:19-64/Domain: amino acid--tRNA ligase repeat homology <ATL>

Query	March	99.3%	Score	2101	DB	1	Length	471											
Best Local	Similarity	99.5%	Pred.	No. 1e-163															
Matches	399	Conservative	0	Mismatches	2	Indels	0	Gaps	0										
QY	1	SNHGDATEAEDEVDVPMVTQTS	SAKGI	YDKLIV	FGSSKID	KELIN	RIERAT	GOREPH	60										
Db	71	SNHGDATEAEDEVDVPMVTQTS	SAKGI	YDKLIV	FGSSKID	KELIN	RIERAT	GOREPH	130										
QY	61	FLRGRIFSSHDMNOVLDAT	ENKKPF	LYTGR	GPSSSEAM	HGHLPI	FFTW	LDVENV	120										
Db	131	FLRGRIFSSHDMNOVLDAT	ENKKPF	LYTGR	GPSSSEAM	HGHLPI	FFTW	LDVENV	190										
QY	121	LVIOMTDEKILMKDLTD	QAYGDA	VENAKDI	IACG	PFINKT	PIFS	DDYGMSSG	GFYKN	180									
Db	191	LVIOMTDEKILMKDLTD	QAYGDA	VENAKDI	IACG	PFINKT	PIFS	DDYGMSSG	GFYKN	250									
QY	181	VVKIQKHVTFNQVKGIF	GFTSD	SGIKIS	FPALQ	APSPFS	FPQIF	FRDRTD	ICLI	PCA	240								
Db	251	VVKIQKHVTFNQVKGIF	GFTSD	SGIKIS	FPALQ	APSPFS	FPQIF	FRDRTD	ICLI	PCA	310								
QY	241	IDDDPYEPMRTDVA	PRIGYPR	KPALHST	FFPALQ	AGQYK	MA	SPNNS	IFLT	DTAKQ	KKT	300							
Db	311	IDDDPYEPMRTDVA	PRIGYPR	KPALHST	FFPALQ	AGQYK	MA	SPNNS	IFLT	DTAKQ	KKT	370							
QY	301	KVNKHA	FGSGRDT	IEEH	QFG	NC	DV	VS	FM	YLF	FE	DDDK	LEQIR	RDY	TSG	AM	TGEL	360	
Db	371	KVNKHA	FGSGRDT	IEEH	QFG	NC	DV	VS	FM	YLF	FE	DDDK	LEQIR	RDY	TSG	AM	TGEL	430	
QY	361	KKALIE	VIQPL	PLAE	HQ	AR	KE	VT	DE	IV	KE	FM	PK	KL	S	DFQ	401		
Db	431	KKALIE	VIQPL	PLAE	HQ	AR	KE	VT	DE	IV	KE	FM	PK	KL	S	DFQ	471		

RESULT 2

TMBO

tryptophan--trna ligase (EC 6.1.1.2) [validated] - bovine

N:Alternate names: tryptophanyl-trna synthetase

C:Species: Bos primigenius taurus (cattle)

C:Date: 30-Sep-1992 #sequence,revision 30-Sep-1992 #text-change 21-Jul-2000

A:Accession: A40279; JN0354; S10460; S14540

R:Garett, M.; Palot, B.; Trezevant, V.; Labouesse, J.; Merle, M.; Gandar, J.C.; Bendeth

Biochemistry 30, 7809-7817, 1991

A:Title: A mammalian tryptophanyl-trna synthetase shows little homology to prokaryotic

A:Reference number: A40279; MUID:91329348

A:Accession: A40279

A:Molecule type: mRNA

A:Residues: 1-475 <GAR>

A:Cross-references: GB:WMT4074; EMBL:X53918; NID:g163798; PIDN:AAA30799.1; PID:g163799

A:Experimental source: pancreas

A:Note: The authors translated the codon CTG for residue 347 as Ala and CAG for residue

A:Note: Part of this sequence was confirmed by protein sequencing

R:Zagorova, T.A.; Kovaleva, I.G.; Favorova, O.O.; Aevina, H.B.; Telezhinskaya, I.H.

Bioorg. Khim. 15, 1307-1311, 1989

A:Title: Amino acid sequence of several peptides of tryptophanyl-trna synthetase from ca

A:Reference number: JN0354; MUID:90211408

A:Accession: JN0354

A:Molecule type: protein

A:Residues: 112-124;282-287, 'N',288, 'E',289-292, 'Q',293-294, 'IR',336-353;423-441,443-445

A:Experimental source: liver

A:Note: this paper is in Russian

C:Superfamily: mammalian tryptophan--trna ligase; amino acid--trna ligase repeat homolog

**C; Keywords:** aminoacyl-tRNA synthetase; ATP; ligase; protein biosynthesis  
F; 24-69/Domain: amino acid--tRNA ligase repeat homology <ATL>

[illegible]

RESULT 3  
 YNRBPR  
 tryptophan--trNA ligase (EC 6.1.1.2) [validated] - rabbit  
 N:Alternate names: tryptophanyl-trNA synthetase  
 C:Species: *Oryctolagus cuniculus* (domestic rabbit)  
 C:date: 30-Sep-1992 #sequence\_revision 13-Feb-1998 #text\_change 26-May-2000  
 C:Accession: A35904; S37396  
 P:Rec. C.C.; Craigen, W.J.; Muzny, D.M.; Harlow, E.; Caskey, C.T.  
 P:Proc. Natl. Acad. Sci. U.S.A. 87, 3508-3512, 1990  
 A:title: Cloning and expression of a mammalian peptide chain release factor with sequ  
 A:Accession: A35904; MUID:90235043  
 A:Reference number: A35904  
 A:Molecule type: mRNA  
 A:Residues: 1-475 <LE>  
 A:Cross-references: GR:M33460  
 R:F.Rlova, L.Y.; Dalphin, M.E.; Justesen, J.; Powell, R.J.; Druegeon, G.; McCaughan, K  
 E:Bio J. 12, 4013-4019, 1993  
 A:title: Mammalian polypeptide chain release factor and tryptophanyl-trNA synthetase  
 A:Reference number: S37396; MUID:94009008  
 A:Accession: S37396  
 A:Molecule type: mRNA  
 A:Residues: 166-177 <PRO>  
 C:Genetics:  
 A:Gene: WRS  
 C:Complex: homodimer [validated, MUID:94009008]  
 C:Function:  
 A:Description: EC 6.1.1.2 [validated, MUID:94009008]; catalyzes the ATP-dependent form  
 A:Note: mammalian WRS (tryptophanyl-trNA synthetases) and eRF (polypeptide chain relea  
 C:Superfamily: mammalian tryptophan--trNA ligase; amino acid--trNA ligase repeat homo  
 C:Keywords: aminocyl-trNA synthetase; ATP; homodimer; ligase; metalloprotein; protein  
 E:174-177/Region: ATP-binding motif (HXGH)

Query Match	92.08;	Score 1946.5;	DB 1;	Length 475;
Best Local Similarity	90.88;	Pred. No. 4.3e-151;		
Matches 364; Conservative	21;	Mismatches 15;	Indels 1;	Gaps 1;

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 24, 2002, 12:49:45 ; Search time 13.479 Seconds  
(without alignments)  
2858.658 Million cell updates/sec

Title: US-09-813-718-10\_COPY\_71\_471

Perfect score: 2116

Sequence: 1 SNHGPDATGEDEFDVDPWTV.....VTDEIVKEFMTPRKLSDFQ 401

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : PIR\_71:\*

1: p1r1:\*

2: p1r2:\*

3: p1r3:\*

4: p1r4:\*

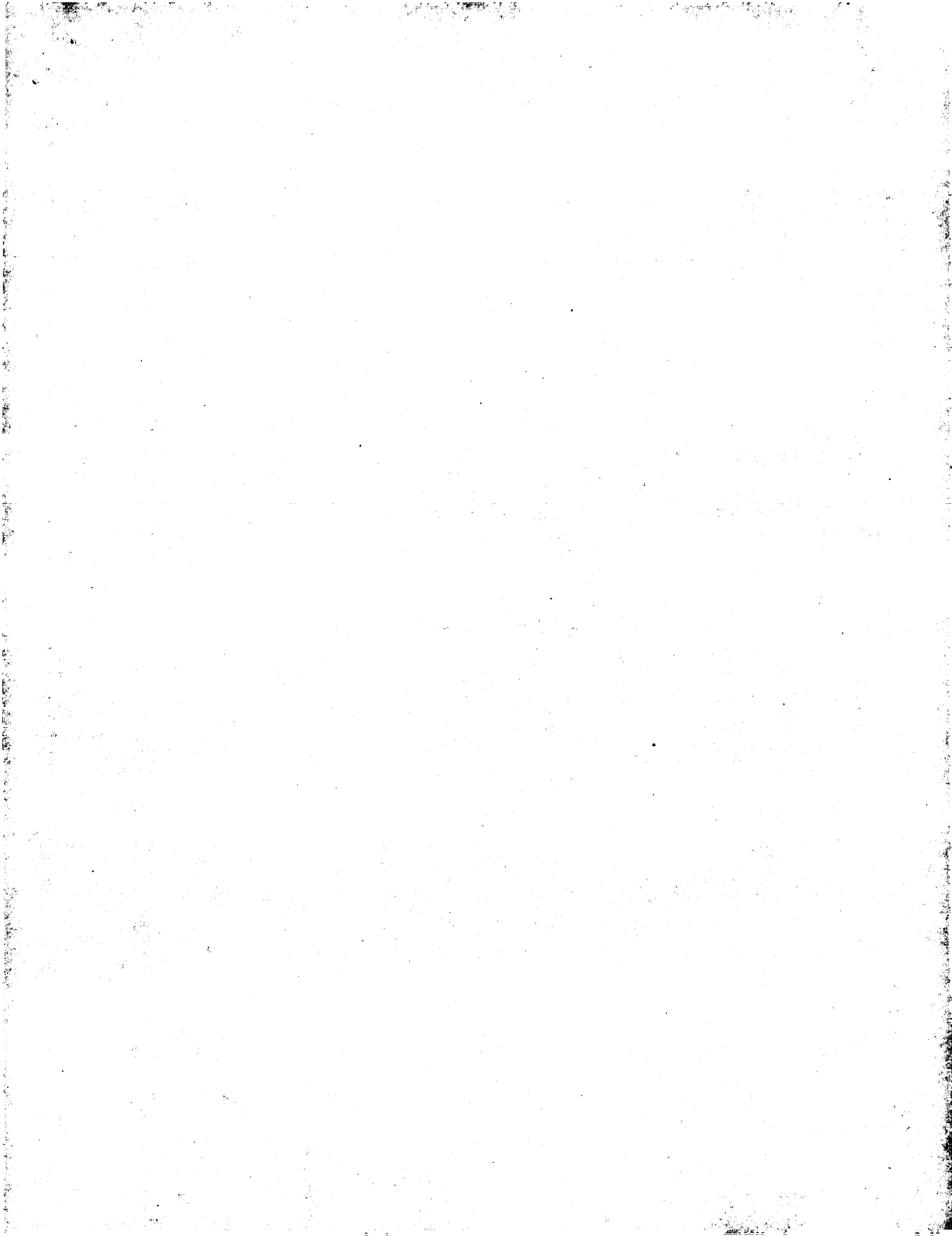
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2101	99.3	471	1 A41706	tyrptophan--trna 1
2	2028.5	95.9	475	1 YWBO	tyrptophan--trna 1
3	1946.5	92.0	475	1 YWRBR	tyrptophan--trna 1
4	1938	91.6	481	2 S50053	tyrptophan--trna 1
5	1210	57.2	395	2 S58157	hypothetical prote
6	1163	55.0	432	2 S51901	tyrptophan--trna 1
7	907	42.9	386	2 C90190	tyrptophanyl--trna
8	803	37.9	385	2 C75020	tyrptophanyl--trna
9	626.5	29.6	301	2 G71206	tyrptophan--trna 1
10	534.5	25.3	380	2 G84373	tyrptophanyl--trna
11	409.5	19.4	370	2 F64476	tyrptophan--trna 1
12	397.5	18.8	364	2 E69131	tyrptophan--trna 1
13	386	18.2	134	2 T43806	tyrptophan--trna 1
14	370.5	17.5	420	2 E69461	tyrptophanyl--trna
15	353	16.7	374	2 D72477	probable tyrptopha
16	269.5	12.7	513	2 F84371	tyrptophanyl--trna
17	192	9.1	323	2 H63346	tyrosyl--trna synth
18	188	8.9	364	2 E72512	probable tyrosyl-t
19	185.5	8.8	341	2 D95260	tyrptophanyl--trna
20	185.5	8.8	341	2 G98125	tyrptophan--trna 1
21	174.5	8.2	341	2 B86633	tyrptophan--trna 1
22	172	8.1	394	2 A45999	tyrosine--trna 1ig
23	170.5	8.1	366	2 S75410	tyrosine--trna 1ig
24	169	8.0	395	2 H70385	tyrptophan--trna 1
25	167	7.9	351	2 E75438	tyrptophanyl--trna
26	166.5	7.9	346	2 B71496	tyrptophan--trna 1
27	164.5	7.8	895	2 A86410	protein F3M18.22
28	163.5	7.7	353	2 E70100	tyrptophan--trna 1
29	163	7.7	460	2 C84750	probable tyrosyl-t

## ALIGNMENTS

RESULT 1  
A41706  
tyrptophan--trna ligase (EC 6.1.1.2) [similarity] - human  
N:Alternate names: Interferon-inducible protein IFP53; peptide-chain release factor h  
C:Species: Homo sapiens (man)  
C>Date: 19-May-2000 #sequence\_revision 19-May-2000 #text\_change 19-May-2000  
C:Accession: A41633; A41706; S19246; JN0676; JN0533; S26287  
R:Flecker, J.; Rasmussen, H.H.; Justesen, J.  
Proc. Natl. Acad. Sci. U.S.A. 88, 11520-11524, 1991  
A:Title: Human interferon gamma potentially induces the synthesis of a 55-kDa protein (9  
A:Reference number: A41633; MUID:92107982  
A:Accession: A41633  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-471 <FILE>  
A:Cross-references: GB:559892; NID:930820; PIDN:CAA42545.1; PID:930821  
R:Rubin, B.Y.; Anderson, S.L.; Xing, L.; Powell, R.J.; Tate, W.P.  
J. Biol. Chem. 266, 24245-24248, 1991  
A:Title: Interferon induces tyrtophanyl--trna synthetase expression in human fibrobla  
A:Reference number: A41706; MUID:92105071  
A:Accession: A41706  
A:Molecule type: mRNA  
A>Status: preliminary  
A:Residues: 1-471 <RUB>  
A:Cross-references: GB:M77804; NID:9184656; PIDN:AAA67324.1; PID:9184657  
R:Buwilt, U.; Flohr, T.; Boettger, E.C.  
EMBO J. 11, 489-496, 1992  
A:Title: Molecular cloning and characterization of an interferon induced human CDNA w  
A:Reference number: S19246; MUID:92164636  
A:Accession: S19246  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-423, 'R', 425-471 <BUW>  
A:Cross-references: EMBL:X62570; NID:932708; PIDN:CAA44450.1; PID:932709  
R:Polova, L.Y.; Grigorjeva, A.Y.; Sudomina, M.A.; Kisselev, L.L.  
Gene 128, 237-245, 1993  
A:Title: The human gene encoding tyrtophanyl--trna synthetase: Interferon-response el  
A:Reference number: JN0676; MUID:93292992  
A:Accession: JN0676  
A:Molecule type: DNA  
A:Residues: 1-141:182-471 <PRO>  
A:Cross-references: GB:X67918; GB:S62837; NID:937968; GB:X67919; NID:937969; GB:X6792  
4; NID:937974; GB:X67925; GB:S62855; NID:937975; GB:X67926; GB:S62856; NID:937976; GB  
A:Note: the authors translated the codon GGC for residue 55 as Cys and GAG for residu  
A:Note: this translation is not annotated in GenBank entries HSRXSXA, HSRXSXB, HSR  
R:Polova, L.Y.; Sudomina, M.A.; Grigorjeva, A.Y.; Zinovjeva, O.L.; Kisselev, L.L.  
Gene 109, 291-296, 1991  
A:Title: Cloning and nucleotide sequence of the structural gene encoding for human tr  
A:Reference number: JN0533; MUID:92112058  
A:Accession: JN0533  
A:Molecule type: mRNA



Fri Oct 25 09:23:11 2002

us-09-813-718-10\_copy\_71\_471.ra

Page 9

Search completed: October 24, 2002, 12:54:34  
Job time : 10.8029 secs

---





TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-928-100-4

Query Match  
Best Local Similarity 5.1%; Score 108; DB 3; Length 197;  
Matches 42; Conservative 27; Mismatches 73; Indels 24; Gaps 6;

QY 237 IPCAIDDDPYFRMTRD-----VAPRIGYKPKPALHSTFFPALOGAQTMSA 282  
DB 5 VPGVTDGKPMIEQTRREIVRSFNNAVNCVLYVEPGIYPENE--RAGRLPGLDG--NAKMSK 61

QY 283 SDPNSSIFLDDTAQKQITKYNKHAFFSGGRDTIEHRQFGNCVDVSEFWLTFE--LEDD 340  
DB 62 S-LNNGIYLDADDATLTKRKWSMTDPDHIRVEDPGKIEGN---WVFHYLDVFGRPEDA 116

QY 341 DKLEQIRKDYTGAMLTGELKALIEVLOPLIAEHQARRKRVETDEI 386  
DB 117 QEIDMKERYQRGGLGVKTKRYLLELRELGPTRERRIEFKAMD 162

RESULT 12  
US-09-183-134-2  
Sequence 2, Application US/09183134  
Patent No. 6165759  
GENERAL INFORMATION:  
APPLICANT: Lawlor, Elizabeth  
TITLE OF INVENTION: NOVEL TRYPTOPHANYL TRNA SYNTHETASE  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Smithline Beecham Corporation  
STREET: 709 Swedeland Road  
CITY: King of Prussia  
STATE: PA  
COUNTRY: USA  
ZIP: 19406-0939  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/183,134  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Gimm1, Edward R  
REGISTRATION NUMBER: 38,891  
REFERENCE/DOCKET NUMBER: P31624  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-270-4478  
TELEFAX: 610-270-5090  
TELEX:  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 197 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-183-134-2

Query Match  
Best Local Similarity 5.1%; Score 108; DB 4; Length 197;  
Matches 42; Conservative 27; Mismatches 73; Indels 24; Gaps 6;

QY 237 IPCAIDDDPYFRMTRD-----VAPRIGYKPKPALHSTFFPALOGAQTMSA 282  
DB 5 VPGVTDGKPMIEQTRREIVRSFNNAVNCVLYVEPGIYPENE--RAGRLPGLDG--NAKMSK 61

DB 5 VPGVTDGKPMIEQTRREIVRSFNNAVNCVLYVEPGIYPENE--RAGRLPGLDG--NAKMSK 61

QY 283 SDPNSSIFLDDTAQKQITKYNKHAFFSGGRDTIEHRQFGNCVDVSEFWLTFE--LEDD 340  
DB 62 S-LNNGIYLDADDATLTKRKWSMTDPDHIRVEDPGKIEGN---WVFHYLDVFGRPEDA 116

QY 341 DKLEQIRKDYTGAMLTGELKALIEVLOPLIAEHQARRKRVETDEI 386  
DB 117 QEIDMKERYQRGGLGVKTKRYLLELRELGPTRERRIEFKAMD 162

RESULT 13  
US-09-492-581-4  
Sequence 4, Application US/09492581  
Patent No. 6346409  
GENERAL INFORMATION:  
APPLICANT: Gentry, Danlle  
APPLICANT: Greenwood, Claire  
APPLICANT: Lawlor, Elizabeth  
TITLE OF INVENTION: No. 6346409el trps  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Smithline Beecham Corporation  
STREET: 709 Swedeland Road  
CITY: King of Prussia  
STATE: PA  
COUNTRY: USA  
ZIP: 19406-0939  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/492,581  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/928,100  
FILING DATE: 12-SEP-1997  
APPLICATION NUMBER: 9619072.3  
FILING DATE: 12-SEP-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Gimm1, Edward R  
REGISTRATION NUMBER: 38,891  
REFERENCE/DOCKET NUMBER: P31624-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-270-4478  
TELEFAX: 610-270-5090  
TELEX:  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 197 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-492-581-4

Query Match  
Best Local Similarity 5.1%; Score 108; DB 4; Length 197;  
Matches 42; Conservative 27; Mismatches 73; Indels 24; Gaps 6;

QY 237 IPCAIDDDPYFRMTRD-----VAPRIGYKPKPALHSTFFPALOGAQTMSA 282  
DB 5 VPGVTDGKPMIEQTRREIVRSFNNAVNCVLYVEPGIYPENE--RAGRLPGLDG--NAKMSK 61

QY 283 SDPNSSIFLDDTAQKQITKYNKHAFFSGGRDTIEHRQFGNCVDVSEFWLTFE--LEDD 340  
DB 62 S-LNNGIYLDADDATLTKRKWSMTDPDHIRVEDPGKIEGN---WVFHYLDVFGRPEDA 116

QY 341 DKLEQIRKDYTGAMLTGELKALIEVLOPLIAEHQARRKRVETDEI 386  
DB 117 QEIDMKERYQRGGLGVKTKRYLLELRELGPTRERRIEFKAMD 162





```

; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 409 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-743-130A-2

Query Match
Best Local Similarity 20.7%; Score 162.5; DB 2; Length 409;
Matches 78; Conservative 59; Mismatches 132; Indels 107; Gaps 16;

QY 76 VLDAVENK-KPEFLYTGSGPSSEAMHGHLPFI-----FTKMLQDV-----F 117
D 27 IKDVLKERNRPVKIYWGAPTKGP-HCGYFVPMIKLHFLKAGCEVTVLADLHAFLLDM 85
QY 118 NPELYIOMTDEKYLKMLTLDDQAYGDAVENAKDIIACGFDINKTFITSDLYMGMSGGF 177
D 86 KAPLEVYKRAKYEEFVVKALKLSINPIERLKFVYGSSYOKGGDYV--MDLEKLSNIV 142
QY 178 YKNVVK-----IOKHVFNOYKGIETGFTDSDICIKISFPALQAAFSFNSFPQIFRDRTD 232
D 143 SQNDARRAGADVKKVANPLISGLT-----TPLMQA-----IDEEHLG 180
QY 233 IOCLIPCAIDDPYFRMTRDVAPRIGYKPKPALHSTFFPALQAGQTKMSASDPNSSIFLT 292
D 181 VDAQRG-GVDQRKIFVLAEMENLPSIGYKKRAHLMNPMVPGI-GGGKMSASDPNSKIDII 238
QY 293 DTAKOIKTKVKNKHAFIGSG--RDT-----IIEHRQFCGN 323
D 239 EEPKVVKKRKSAYCAPGELKDNGLIAFLIEVYIOPIAELKTGVEGAFKLDIDRPERKYG- 297
QY 324 CDVDSFMYLTFLEDDDKLEQIRKDYTSGAMLTGELKKALL-----EVLQPLIAEQARR 379
D 298 ---PLST-----DSIQDLKADVDKGLAPPLKLVADKINELLAPLRAEFESS- 343
QY 380 KEVTDEIVKEFMTPRK 395
D 344 -----EEFQVAAK 351

RESULT 6
US-08-705-868-4
; Sequence 4, Application US/08705868
; Patent No. 5885798
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Coleman, Roger
; APPLICANT: Au-Young, Janice
; APPLICANT: Murry, Lynn E.
; TITLE OF INVENTION: NOVEL MONOCYTE ACTIVATING CYTOKINE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: U.S.
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/705,868
; FILING DATE: Filed Herewith
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0117 US
; TELECOMMUNICATION INFORMATION:

```

```

; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 388 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 1184699
US-08-705-868-4

Query Match
Best Local Similarity 21.2%; Score 132.5; DB 2; Length 388;
Matches 79; Conservative 64; Mismatches 149; Indels 80; Gaps 19;

QY 60 HFLRRGI--FESHDMQVLDAYENKKPEFLDYTGSGPSSEAMHGHLPFI--FTKMLQDV 116
D 12 HLIFRNLOEVGEEKLKEILKERELK---LYWGATATGKP-HVAYFVPMKIDFLKAG 66
QY 117 FNVPLVT---QMTDEKYLKDLTLDDQAYGDAVENAKDIIACGFDINKTFITSDLYMG 172
D 67 CEVTLLEFADLHAYLDNMKAPWELLELRVSYEYENYIKAM-LESIGVLEKLFITGTDYQ- 124
QY 173 MSSGFEKVVVIOKHVFNOYK---GIEGFTDSDICIKISFPALQAAFSFNSFPQIFR 228
D 125 LSKETTLDYRLSSVYVTHDSKKAGAEVYKOVENHLSGLIYPLGIA---LDEEYLYK-- 179
QY 229 DRTDIQCLIPCAIDDPYFRMTRDVAPRIGYKPKPALHSTFFPALQAGQTKMSASDPNSS 288
D 180 ---DAQF---GIDQRKIFTFAEKYLPAIGYSKRVHLMNPMVPGITG--SKMSSEESK 231
QY 289 IFLVDTAQOIKTKVKNKAF-----SGGRDTIEH-----KQFGNCDVD 327
D 232 IDLDRKEDVKKLKLK-AFCERGVNENNGVLSFKHVLFPKSEFVYLDRKMGCN--- 286
QY 328 VSFMYLTFLEDDDKLEQIRKDYTSGAMLTGELKKALLIEVLOPLIAEQARRKEVTDIV 387
D 287 ---KTYTAYVD-----LEKDFAAEVVHPGDLKNSVVALNKL-----DPIR 325
QY 388 KEFMTPR--RKLS 397
D 326 EKENTPALKILA 337

RESULT 7
US-09-123-615-4
; Sequence 4, Application US/09123615
; Patent No. 6090377
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Coleman, Roger
; APPLICANT: Au-Young, Janice
; APPLICANT: Murry, Lynn E.
; TITLE OF INVENTION: NOVEL MONOCYTE ACTIVATING CYTOKINE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: U.S.
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/123,615
; FILING DATE:

```

US-09-492-581-2

Query Match	8.88;	Score 185.5;	DB 4;	Length 341;
Best Local Similarity	24.08;	Pred. No. 2.8e-12;		
Matches	81;	Conservative	60;	Mismatches 128;
			Indels	69;
			Gaps	15;

QY	84	KPEFLYRGSSSAMHGHLEPIEFKRLMODVENVPLMDTMDXK--LWKDLTLLOAV	142
Dd	3	KPIIL-TGDRTGK-LHGHVYVSLKRR-----VLGRBEDIKYMDFVLADQAL	49
QY	143	GDAVEN-----AKDIACGFIDNKTFF--SDLDYMGSSGFYKNV---KIQK	186
Dd	50	TDAHKDPOTIVESIGNVALDYLADGLDPNKSITPISQIDPELALSMYNNVLSPARLER	109
QY	187	HAFENOVKGIQFTDSDICIKISEPAAIOA--PSFSNSPQIFRDNRDIOCLIPCARIDQ	244
Dd	110	NPTKTELSQKGFESIPITGFLVYPIAQAADITAFKANY-----VPVGTDK	156
QY	245	PYFKMTD-----VAPRIQYRKPALHSTFPFALOGAOTKMSASDPNSSIF	290
Dd	157	PMIEOTREIYVFSFNNAAYNCVDLYVEEGYLPENE--RAGRLEPGIDG--NAKMSK-S--LNGIY	212
QY	291	LTFDAKQIKETVKNHAFSGGRDTEEHROFGNCDVNVSEMYLTF--LEDDKLEQIRK	348
Dd	213	LADADLTLRKVMYMDPDHRIREDPDGKIEGN---MVFHYLDVFORPDADQIAOMKE	266
QY	349	DYTSGAMLTGELKALLEVLOPLAEHQARKEVTDDEI	386
Dd	269	RYORGGLGVATKRYLLEILEREGLPIRREKTEFAKAM	306

## RESULT 4

US-08-743-130A-39  
; Sequence 39, Application US/08743130A

```

? GENERAL INFORMATION:
? APPLICANT: Sassanfar, Mandana
? APPLICANT: Gallant, Paul L.
? APPLICANT: Shen, Xiaoyu
? APPLICANT: Tao, Mianjun
? APPLICANT: Tao, Jianshi
? APPLICANT: Houman, Fariba
? TITLE OF INVENTION: CANDIDA TYROSILO-tRNA SYNTHETASE
? TITLE OF INVENTION: PROTEINS, NUCLEIC ACIDS AND STRAINS COMPRISING SAME
? NUMBER OF SEQUENCES: 41
? CORRESPONDENCE ADDRESS:

```

```

:
: COMPUTER READABLE FORM:
:
: MEDIUM TYPE: Floppy disk
:
: COMPUTER: IBM PC compatible
:
: OPERATING SYSTEM: PC-DOS/MS-DOS
:
: SOFTWARE: PatentIn Release #1.0, Version #1.30A
:
: CURRENT APPLICATION DATA:
:
: APPLICATION NUMBER: US/08/743,130A
:

```

US-08-743-130A-39

Query Match	7.8%;	Score 164.5;	DB 2;	Length 409;
Best Local Similarity	20.7%;	Pred. No. 8.8e-10;		
Matches	78;	Conservative	59;	Mismatches 132;
			Indels	107;
			Gaps	16;

```

QY 76 VLDVYENK-KPEVLYUTGSGPSEAMHYGHILPFI-----FTKMLQDV-----F 117
Db 27 IKDVLKEKNRPVKYITWGSTARPGKR-HCGYFVPMKLAHFLKAGCEVYLLADLHAFLADNM 85
QY 118 NVPLVIQMTDDEKYLMDLTLDAQYGADEVANAKDILACGEDINKFTFESDLDYGMSSGF 177
Db 86 KAPLEVVYKRAKYYEFVYKAILKINSINPIELKTKVVGSSYQKGGDY---MDLFKLSNIY 142
QY 178 YKKNVYK-----IOKHVTFRNOVKKGIFFGTDSDCIGKISFPALQAAFSFNSGFPIERDRPTD 233
Db 143 SQNDAKRKAGADVYQKVAANPLLSGLI-----YPLMQA-----IDEHILG 180
QY 233 IQCLIPCALIDDPYFRMRIDVAPRPIGYPKRALLSHTFFPALOGQOTKMSADPNSSIFLT 292
Db 181 VDAQFG-GVDDRCIKFEVLAEENLPSIGYKRAHLNMPVPGI-GGGKMSASDPNSKIDII 238
QY 293 DTAOKIITKYVKNHAFSGS--RDT-----IEHRPGCN 323
Db 239 EEPKYVKKKNVSAATCAQELKDNOLIAFIEVYIOPIALKTVGEGAFKDIIDREKKGG- 297
QY 324 CDVDSFMYLTFLEDDDKLEIIRKDYTSGLMTGLKALI---EVLQPLIAEHQARR 379
Db 298 ---PLSY-----DSIEQLKADFDVCKLAPPDLSGVADKINELLAPIRAEFESS- 343
QY 380 KEVYDELYKEFMTPRK 395
Db 344 -----EEFOVAOK 351

```

## RESULT 5

US-08-743-130A-2  
; Sequence 2, Application US/08743130A

```

GENERAL INFORMATION:
APPLICANT: Sassanfar, Mandana
APPLICANT: Gallant, Paul L.
APPLICANT: Shen, Xiaoyu
APPLICANT: Tao, Nianjun
APPLICANT: Tao, Jianshi
APPLICANT: Houman, Fariba
TITLE OF INVENTION: CANDIDA TYROSYL-tRNA SYNTHETASE
TITLE OF INVENTION: PROTEINS, NUCLEIC ACIDS AND STRAINS COMPRISING SAME
NUMBER OF SEQUENCES: 41
CORRESPONDENCE ADDRESS:

```

```

1: COMPUTER READABLE FORM:
2: MEDIUM TYPE: Floppy disk
3: COMPUTER: IBM PC compatible
4: OPERATING SYSTEM: PC-DOS/MS-DOS
5: SOFTWARE: Patientin Release #1.0, Version #1.30
6: CURRENT APPLICATION DATA:
7: APPLICATION NUMBER: US/08/743,130A
8: FILING DATE: 01-NOV-1986

```

```

: INFORMATION FOR SEQ ID NO: 39:
:
: SEQUENCE CHARACTERISTICS:
:
: LENGTH: 409 amino acids
:
: TYPE: amino acid
:
: STRANDEDNESS:
:

```

```

1      RESULT 2
2      US-08-928-100-2
3      : Sequence 2, Application US/08928100
4      : Patent No. 6046174
5      :
6      : GENERAL INFORMATION:
7      : APPLICANT: Gentry, Danile
8      : APPLICANT: Greenwood, Claire
9      : APPLICANT: Lawlor, Elizabeth
10     : TITLE OF INVENTION: No. 6046174el tips
11     :
12     : NUMBER OF SEQUENCES: 6
13     :
14     : CORRESPONDENCE ADDRESS:
15     : ADDRESSEE: SmltnKline Beecham Corporation
16     : STREET: 709 Swedeland Road
17     : CITY: King of Prussia
18     : STATE: PA
19     :
20     : COUNTRY: USA
21     : ZIP: 19406-0939
22     :
23     : COMPUTER READABLE FORM:
24     : MEDIUM TYPE: Diskette
25     : COMPUTER: IBM Compatible
26     : OPERATING SYSTEM: DOS
27     :
28     : SOFTWARE: FastSeq for Windows Version 2.0
29     :
30     : CURRENT APPLICATION DATA:
31     : APPLICATION NUMBER: US/08/928,100
32     : FILING DATE: 12-SEP-1997
33     : CLASSIFICATION: 424
34     :
35     : PRIOR APPLICATION DATA:
36     : APPLICATION NUMBER: 9619072.3
37     : FILING DATE: 12-SEP-1996
38     :
39     : ATTORNEY/AGENT INFORMATION:
40     : NAME: Gimm1, Edward R
41     : REGISTRATION NUMBER: 38,891
42     : REFERENCE/DOCKET NUMBER: P31624-1
43     :
44     : TELECOMMUNICATION INFORMATION:
45     : TELEPHONE: 610-270-4478
46     : TELEFAX: 610-270-5090
47     :
48     : TELEX:
49     :
50     : INFORMATION FOR SEQ ID NO: 2:
51     :
52     : SEQUENCE CHARACTERISTICS:
53     : LENGTH: 341 amino acids
54     : TYPE: amino acid
55     : STRANDEDNESS: single
56     : TOPOLOGY: linear
57     :
58     : MOLECULE TYPE: protein
59     :
60     : US-08-928-100-2
61     :
62     : Query Match 8.88; Score 185.5; DB 3; Length 341;
63     : Best Local Similarity 24.08; Pred. No. 2.be-12;

```

```

1      RESULT 3
2      US-09-492-581-2
3      ; Sequence 2, Application US/09492581
4      ; Patent No. 6346409
5      ; GENERAL INFORMATION:
6      ; APPLICANT: Gentry, Danlle
7      ; APPLICANT: Greenwood, Claire
8      ; APPLICANT: Lawlor, Elizabeth
9      ; TITLE OF INVENTION: No. 6346409el trps
10     ; NUMBER OF SEQUENCES: 6
11     ; CORRESPONDENCE ADDRESS:
12     ; ADDRESSEE: SmithKline Beecham Corporation
13     ; STREET: 709 Swedeland Road
14     ; CITY: King of Prussia
15     ; STATE: PA
16     ; COUNTRY: USA
17     ; ZIP: 19406-0939
18     ; COMPUTER READABLE FORM:
19     ; MEDIUM TYPE: Diskette
20     ; COMPUTER: IBM Compatible
21     ; OPERATING SYSTEM: DOS
22     ; SOFTWARE: FASTSEQ for Windows Version 2.0
23     ; CURRENT APPLICATION DATA:
24     ; APPLICATION NUMBER: US/09/492,581
25     ; FILING DATE:
26     ; CLASSIFICATION:
27     ; PRIOR APPLICATION DATA:
28     ; APPLICATION NUMBER: US/08/928,100
29     ; FILING DATE: 12-SEP-1997
30     ; APPLICATION NUMBER: 9619072.3
31     ; FILING DATE: 12-SEP-1996
32     ; ATTORNEY/AGENT INFORMATION:
33     ; NAME: Gimm1, Edward R
34     ; REGISTRATION NUMBER: 38,891
35     ; REFERENCE/DOCKET NUMBER: P31624-1
36     ; TELECOMMUNICATION INFORMATION:
37     ; TELEPHONE: 610-270-4478
38     ; TELEFAX: 610-270-5090
39     ; TELEX:
40     ; INFORMATION FOR SEQ ID NO: 2:
41     ; SEQUENCE CHARACTERISTICS:
42     ; LENGTH: 341 amino acids
43     ; TYPE: amino acid
44     ; STRANDEDNESS: single
45     ; TOPOLOGY: linear
46     ; MOLECULE TYPE: protein
47     ;

```

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: October 24, 2002, 12:49:45 : Search time 9.8029 Seconds  
(without alignments)  
999.159 Million cell updates/sec

Title: US-09-813-718-10\_COPY\_71\_471

Perfect score: 2116  
Sequence: 1 SNHGPDATAEDEFDVPMWT.....VTDEIVKEMPRKLSDFDQ 401

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

## Database :

Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/1/1aa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/1/1aa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/PCtUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/backfilest1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1218.5	57.6	424	4	US-08-876-885-26 Sequence 26, Appli
2	185.5	8.8	341	3	US-08-928-100-2 Sequence 2, Appli
3	185.5	8.8	341	4	US-09-492-581-2 Sequence 2, Appli
4	164.5	7.8	409	2	US-08-743-130A-39 Sequence 39, Appli
5	162.5	7.7	409	2	US-08-743-130A-2 Sequence 2, Appli
6	132.5	6.3	388	2	US-08-705-868-4 Sequence 4, Appli
7	132.5	6.3	388	3	US-09-123-615-4 Sequence 4, Appli
8	113.5	5.4	418	4	US-08-855-910-11 Sequence 11, Appli
9	109	5.2	377	4	US-09-352-990-28 Sequence 28, Appli
10	108	5.1	197	3	US-08-923-867-2 Sequence 2, Appli
11	108	5.1	197	3	US-08-928-100-4 Sequence 4, Appli
12	108	5.1	197	4	US-09-183-134-2 Sequence 2, Appli
13	108	5.1	197	4	US-09-492-581-4 Sequence 4, Appli
14	104.5	4.9	370	2	US-08-415-593-45 Sequence 45, Appli
15	95	4.5	418	4	US-08-844-054-2 Sequence 2, Appli
16	95	4.5	418	4	US-09-347-333-2 Sequence 8, Appli
17	94.5	4.5	877	2	US-08-907-166-8 Sequence 2, Appli
18	89.5	4.2	344	4	US-09-393-554-2 Sequence 2, Appli
19	88.5	4.2	898	1	US-08-465-995A-4 Sequence 4, Appli
20	88.5	4.2	898	1	US-08-465-994C-4 Sequence 4, Appli
21	88.5	4.2	898	2	US-08-966-145-4 Sequence 4, Appli
22	88.5	4.2	920	1	US-08-101-593-4 Sequence 4, Appli
23	88	4.2	606	2	US-08-883-534-3 Sequence 3, Appli
24	88	4.2	606	3	US-09-204-764-3 Sequence 3, Appli
25	87.5	4.1	855	4	US-08-890-865A-10 Sequence 10, Appli
26	87.5	4.1	898	1	US-08-465-995A-2 Sequence 2, Appli
27	87.5	4.1	898	2	US-08-465-994C-2 Sequence 2, Appli

28	87.5	4.1	898	2	US-08-966-145-2 Sequence 2, Appli
29	87.5	4.1	920	1	US-08-101-593-2 Sequence 2, Appli
30	85	4.0	410	4	US-09-352-990-16 Sequence 16, Appli
31	83.5	3.9	428	3	US-08-331-625A-43 Sequence 43, Appli
32	83.5	3.9	510	1	US-08-249-112-3 Sequence 3, Appli
33	83.5	3.9	510	5	PCT-US95-06556-3 Sequence 7, Appli
34	83.5	3.9	970	1	US-08-375-709-7 Sequence 7, Appli
35	83.5	3.9	970	1	US-08-752-929-7 Sequence 7, Appli
36	83.5	3.9	970	4	US-09-090-793-5 Sequence 5, Appli
37	83.5	3.9	1452	4	US-08-331-625A-2 Sequence 2, Appli
38	83.5	3.9	1452	3	PCT-US93-04384-18 Sequence 18, Appli
39	83.5	3.9	1452	5	PCT-US93-04692-2 Sequence 2, Appli
40	83	3.9	400	4	US-09-352-990-18 Sequence 18, Appli
41	83	3.9	551	1	US-08-484-493-15 Sequence 15, Appli
42	83	3.9	551	1	US-08-484-494-15 Sequence 15, Appli
43	83	3.9	551	2	US-08-345-212-15 Sequence 15, Appli
44	83	3.9	551	4	US-09-249-003-15 Sequence 15, Appli
45	83	3.9	1073	1	US-07-695-564-1 Sequence 1, Appli

## ALIGNMENTS

```
RESULT 1
US-08-876-885-26
: Sequence 26, Application US/08876885
: Patent No. 6174713
:
: GENERAL INFORMATION:
: APPLICANT: Shen, Xiaoyu
: APPLICANT: Houman, Fariba
: TITLE OF INVENTION: CANDIDA CYTOPLASMIC TRYPTOPANYL-TRNA
: TITLE OF INVENTION: SYNTHETASE PROTEINS, NUCLEIC ACIDS AND STRAINS COMPRISING
: TITLE OF INVENTION: SAME
: NUMBER OF SEQUENCES: 26
: CORRESPONDENCE ADDRESS:
: ADDRESS: Hamilton, Brook, Smith & Reynolds, P.C.
: STREET: Two Militia Drive
: CITY: Lexington
: STATE: MA
: COUNTRY: USA
: ZIP: 02173
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/876, 885
: FILING DATE: 16-JUN-1997
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Brook, David E.
: REGISTRATION NUMBER: 22,592
: REFERENCE/DOCKET NUMBER: CP197-02
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (781) 861-6240
: TELEFAX: (781) 861-9540
: INFORMATION FOR SEQ ID NO: 26:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 424 amino acids
: TYPE: amino acid
: STRANDEDNESS:
: TOPOLOGY: linear
: US-08-876-885-26
```

Query Match 57.6%, Score 1218.5, DB 4, Length 424;  
Best Local Similarity 57.1%, Pred. No. 3.4e-128;  
Matches 226, Conservative 71, Mismatches 92, Indels 7, Gaps 3;

OY 8 TEAEEDFVDPWTVQ---TSSAKGIDYDKLIVRFGSSKIDKELNIRIRATGQRPHFRLR 63  
DB 13 TEESQKITPWEVGAVVDGKSMGIDYDKLSQFCTKITTEBTLRFYQVGTGEHPFLK 72







Db 239 KEYGGRMLTGEVKKRLTEVLIEYKRRRAAAVTDEWDAFMAVRPLPSKFE 292

## RESULT 13

AAB58517  
ID AAB58517 standard; Protein; 173 AA.

AC AAB58517;

DT 14-MAR-2001 (first entry)

DE Lung cancer associated polypeptide sequence SEQ ID 855.

XX Human: lung cancer associated protein; neuroprotective; cytostatic;  
KW cardiactive; immunomodulatory; muscular active; vulnerary;  
KW gastrointestinal; nephrotoxic; antineflective; gynecological;  
KW antibacterial; diagnosis; neural disorder; immune disorder; reproductive;  
XX proliferative disorder; wound healing; infectious disease.

OS Homo sapiens.

PN W0200055180-A2.

PD 21-SEP-2000.

PF 08-MAR-2000; 2000WO-US05918.

PR 12-MAR-1999; 99US-0124270.

PA (HUMA-) HUMAN GENOME SCI INC.  
(ROSE/) ROSEN C A.

PI Ruben SM;

XX WPI: 2000-587514/55.

DR N-PSDB; AAF18393.

PT Lung cancer associated gene sequences, referred to as lung cancer  
PT antigens, useful for treatment, prevention, and diagnosis of disorders  
PT such as lung cancer -

XX Claim 11; Page 1392-1393; 1425p; English.

XX Polynucleotide sequences AAF17982 - AAF18424 encode human lung cancer  
CC associated proteins represented in AAB58106 - AAB58548. Lung cancer  
CC associated proteins and polynucleotide sequences, their agonists, and  
CC antagonists may have neuroprotective; cytostatic; cardiactive;  
CC immunomodulatory; muscular active general; vulnerary; gastrointestinal  
CC general; nephrotoxic; antineflective; gynecological; or antibacterial  
CC activity. The invention also includes antibodies specific for the  
CC protein or polynucleotide sequences. The lung cancer associated  
CC polynucleotide sequences may be used for detection of lung cancer,  
CC chromosome identification, as chromosome markers, and for numerous other  
CC diagnostic or research purposes. The proteins may be used to treat  
CC disorders such as neural, immune, muscular, reproductive,  
CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative  
CC disorders. The proteins may also be used in the treatment of wounds and  
CC infectious diseases. Polynucleotide sequences AAF18425 - AAF18433 and  
CC peptide AAB58549 are used in the course of the invention for the  
CC identification and characterisation of the polynucleotide and protein  
CC sequences.

XX Sequence 173 AA;

XX Query Match 39.3%; Score 831; DB 21; Length 173;

XX Best Local Similarity 98.7%; Pred. No. 1.9e-77;

XX Matches 156; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 107 FLETKMLDVFVNPVLIQMTDEKYLKMDLTLDQAYGDAVENAKDIACGPDINKTFIFS 166  
Db 1 FLETKMLDVFVNPVLIQMTDEKYLKMDLTLDQAYGDAVENAKDIACGPDINKTFIFS 60

Qy 167 DLDYMGSSGFYKNNVYKIQKHVTFNQVKGIFGFTDSDICIGKISFPALQAAFSNSFPQI 226  
Db 61 DLDYMGSSGFYKNNVYKIQKHVTFNQVKGIFGFTDSDICIGKISFPALQAAFSNSFPQI 120

Qy 227 FRDRTDIOCLIPCALIDODPYFRMTDVAAPRIGYPPAL 264

Db 121 FRDRTDIOCLIPCALIDODPYFRMTDVAAPRIGYPPAL 158

## RESULT 14

AAB96409  
ID AAB96409 standard; Protein; 385 AA.

AC AAB96409;

DT 29-OCT-2001 (first entry)

DE Putative P. abyssi tryptophanyl-tRNA synthetase.

XX Hyperthermophilic archaeon; hyperthermophilic protein.

XX Pyrococcus abyssi.

OS Pyrococcus abyssi.

PN FR2792651-A1.

PD 27-OCT-2000.

PF 21-APR-1999; 99FR-0005034.

PR 21-APR-1999; 99FR-0005034.

PA (CNRS ) CNRS CENT NAT RECH SCI.  
(IFRE-) IFREMER INST FR RECH EXPL MER.

PI Forterre P, Thierry JC, Prieur D, Dietrich J, Lecompte O;  
PI Querellou J, Weissenbach J, Saurin W, Heilig R;

DR WPI: 2001-126236/14.

XX New nucleotide sequences isolated from Pyrococcus abyssi encode  
PT proteins useful in industry -

XX Claim 7; Pages 1099-1100; 1657p; French.

XX The present invention relates to the genomic sequence of Pyrococcus  
CC abyssi (see AAF6431 and AAH41223-7) and P. abyssi proteins. P. abyssi is  
CC a hyperthermophilic archaeon, which is isolated from deep-sea  
CC hydrothermal vents. The present sequence is one such P. abyssi protein.  
CC The proteins of the present invention have various potential industrial  
CC uses, since the proteins are stable at very high temperatures, some up to  
CC 110 degrees centigrade.  
CC Note: This patent is in the same patent family as W0200065062, which  
CC contains additional sequences as shown in AAB99132-AAB99143,  
CC AAH75903-AAH75920 and AAG66436.

XX Sequence 385 AA;

XX Query Match 37.9%; Score 803; DB 22; Length 385;

XX Best Local Similarity 45.3%; Pred. No. 4.9e-74;

XX Matches 173; Conservative 63; Mismatches 120; Indels 26; Gaps 9;

Qy 12 EDF-VDPMTVQTSKAGIDYKLIIVFGSSKIDKLIRINRIEATGGRHFLRNGIFPSH 70

Db 3 EDFKTVPEWEGV---VDYKNLIRHFGTSPTEBELKETALETSELPLDFRRKFFPSH 58

Qy 71 RDMANVDAYENKRPFYLYTGRGSSAMHVGHLIPFTKMLDVFVNPVLIQMTDEK 130

Db 59 RDKYVLDYDEGRGFLYTGGRGSG-PMHIGHILIPFATKMLQKFGVNLVYIQITDEK 117

Qy 131 YLWMD-LFLDQAYGDAVENAKDIACGPDINKTFIFSDLDYMGSSGFYKNNVYKIQHVT 189

Db 118 FLKENTLFEDTKHWAYENILDIIVAGFPDPTFLFQNSF---TKIYEMAIPIAKKIN 173

PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139889.  
PR 23-JUN-1999; 99US-0140353.  
PR 24-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149358.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.

PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160989.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 45.7%; Score 968; DB 21; Length 292;  
Best Local Similarity 60.9%; Pred. No. 2,9e-91;  
Matches 179; Conservative 51; Mismatches 62; Indels 2; Gaps 1;

Qy 108 IFTKWLQDVFNVPVLTIOGTDEKYLKMDLTLDQAGDVENAKDITACGPIKTFIRSD 167  
Db 1 MFTKLOEAFKVPVLTIOGLTDDKSTKMLSVESQRLRENNKDDITACGFDVTKTFIRSD 60  
Qy 168 LDYMGMSGFYKNVYKIOKHVTFNOVKCIFGFTSDCIGKISFPALQAPFSNSFPOLF 227  
Db 61 FDYVC--GAFYKNVYKVCVTLNKAMGIFGSGSDPIAKLSFPVQAVPFPSSFPILF 118  
Qy 228 RDRFDIOCLPCALDQDQYFPMTRDVAPRIGYKPKALHSTFFPALQCAQTKMSADPNS 287  
Db 119 PGKDNLRCLPCALDQDQYFPMTRDVAPRLGYSKPALIESTFFPALQGENKMSADPNS 178  
Qy 288 SIFLDTFAKOIKTYKNKAFSGGRDTTEHROFGNCDQVDSFWYLFELDDDKLEDR 347  
Db 179 AIYVTSKADIKRINKIRYAFSGGQDSIKHRELGNLEVDIPVKLSFLEDDSELEHIK 238  
Qy 348 KDYSGAMLTELKALIEVLQPLAEHQARRKEVTDEIVKEFMTPRKLSFDFQ 401



PR	14-JUL-1999;	9905-0144624
PR	15-JUL-1999;	9905-0144605
PR	16-JUL-1999;	9905-0144085
PR	16-JUL-1999;	9905-0144326
PR	19-JUL-1999;	9905-0144305
PR	19-JUL-1999;	9905-0144331
PR	19-JUL-1999;	9905-0144332
PR	19-JUL-1999;	9905-0144333
PR	19-JUL-1999;	9905-0144334
PR	19-JUL-1999;	9905-0144335
PR	20-JUL-1999;	9905-0144352
PR	20-JUL-1999;	9905-0144632
PR	20-JUL-1999;	9905-0144684
PR	21-JUL-1999;	9905-0144814
PR	21-JUL-1999;	9905-0145086
PR	21-JUL-1999;	9905-0145218
PR	23-JUL-1999;	9905-0145315
PR	23-JUL-1999;	9905-0145085
PR	23-JUL-1999;	9905-0145224
PR	26-JUL-1999;	9905-0145427
PR	27-JUL-1999;	9905-0145389
PR	27-JUL-1999;	9905-0145913
PR	27-JUL-1999;	9905-0145919
PR	28-JUL-1999;	9905-0145951
PR	02-AUG-1999;	9905-0146386
PR	02-AUG-1999;	9905-0146388
PR	02-AUG-1999;	9905-0146389
PR	03-AUG-1999;	9905-0147038
PR	06-AUG-1999;	9905-0147416
PR	06-AUG-1999;	9905-0147433
PR	05-AUG-1999;	9905-0147126
PR	05-AUG-1999;	9905-0147192
PR	05-AUG-1999;	9905-0147302
PR	04-AUG-1999;	9905-0147204
PR	04-AUG-1999;	9905-0147430
PR	09-AUG-1999;	9905-0147493
PR	09-AUG-1999;	9905-0147935
PR	10-AUG-1999;	9905-0148171
PR	11-AUG-1999;	9905-0148341
PR	12-AUG-1999;	9905-0148341
PR	13-AUG-1999;	9905-0148365
PR	13-AUG-1999;	9905-0148368
PR	16-AUG-1999;	9905-0149372
PR	17-AUG-1999;	9905-0149175
PR	18-AUG-1999;	9905-0149416
PR	20-AUG-1999;	9905-0148722
PR	20-AUG-1999;	9905-0149723
PR	20-AUG-1999;	9905-0149729
PR	23-AUG-1999;	9905-0149902
PR	23-AUG-1999;	9905-0149930
PR	25-AUG-1999;	9905-0150564
PR	26-AUG-1999;	9905-0150884
PR	27-AUG-1999;	9905-0151065
PR	27-AUG-1999;	9905-0151066
PR	27-AUG-1999;	9905-0151080
PR	30-AUG-1999;	9905-0151303
PR	31-AUG-1999;	9905-0151438
PR	01-SEP-1999;	9905-0151930
PR	07-SEP-1999;	9905-0152363
PR	07-SEP-1999;	9905-0153707
PR	13-SEP-1999;	9905-0153758
PR	15-SEP-1999;	9905-0154018
PR	16-SEP-1999;	9905-0154039
PR	20-SEP-1999;	9905-0154779
PR	22-SEP-1999;	9905-0155139
PR	23-SEP-1999;	9905-0155486
PR	24-SEP-1999;	9905-0155569
PR	28-SEP-1999;	9905-0156458
PR	29-SEP-1999;	9905-0156596
PR	04-OCT-1999;	9905-0157117

```
PR 21-OCT-1999; 9905-0160770.
PR 21-OCT-1999; 9905-0160814.
PR 21-OCT-1999; 9905-0160815.
PR 22-OCT-1999; 9905-0160980.
PR 22-OCT-1999; 9905-0160981.
PR 22-OCT-1999; 9905-0160989.
PR 25-OCT-1999; 9905-0161404.
PR 25-OCT-1999; 9905-0161405.
PR 25-OCT-1999; 9905-0161406.
PR 26-OCT-1999; 9905-0161359.
PR 26-OCT-1999; 9905-0161360.
PR 26-OCT-1999; 9905-0161361.
PR 28-OCT-1999; 9905-0161920.
PR 28-OCT-1999; 9905-0161992.
PR 28-OCT-1999; 9905-0161993.
PR 29-OCT-1999; 9905-0162142.

Query Match 61.6%; Score 1304; DB 21; Length 402;
Best Local Similarity 60.3%; Pred. No. 9,5e-126;
Matches 240; Conservative 67; Mismatches 87; Indels 4; Gaps 2;

QY 6 DATEAE--EDVVDPTVOTSSAKGIDYKLYREGSSKIDKELNIRATGQPHHFLR 63
   | ||| | :||| | ||||| :|| :||:| | :|| |||
Db 7 DERAESESOYVNPWEVSADKDGKIDYDKLIDKFCQRLDSESLIDRVQRLTSROPHVELR 66
   | :||:||||| :||| | ||||| ||||| :|||:| | :||| |
QY 64 RGIFESHDMNOVLDAFYENKKPFYLYTGRGPPSEAMHGHILPIFTKWLQDVFNVLVI 123
   | :||:||||| :||| | ||||| ||||| :|||:| | :||| |
Db 67 RSVFPAHDFNEILDAYRGDKFYLYTGRGPPSEALHGHILPEFTYKLDQAFVPLVI 126
   | :||:||||| :||| | ||||| ||||| :|||:| | :||| |
QY 124 QMTDEKYLKMDLTDQAYGDAVENAKDIACGFEDINKFTFSDIDYGMSSGFYKNVYK 183
   | :||| | :||:| | :||| | ||||| :||| | :||| |
Db 127 QLTDEKSIKKNLSSEESQRLARENAKDIIACGFVYTKTFISDRDYVG--GAFYKNVYK 184
   | :||| | :||:| | :||| | ||||| :||| | :||| |
QY 184 IOKHTEFNQVNGIGFEFTSDCGIKISFPALQAPSESNFSQIREDRDICLIPCALDQ 243
   | :||| | :||:| | :||| | ||||| :||| | :||| |
Db 185 VGKCVTLTKAMGIFGFESEEDPLAKISFPVQAVSFSPFHLFGKDKNLKCLIPCALDQ 244
   | :||| | :||:| | :||| | ||||| :||| | :||| |
QY 244 DPFYMRDVAAPRIGYPRPALHSTFPALOGAOTKMSASDPNSSIFLDTAKQIKTFVN 303
   | :||| | :||:| | :||| | ||||| :||| | :||| |
Db 245 DPFYMRDVAAPRIGYSPALIESTFPPALOGENGKMSASDPNSAIYVTSKAKDKINKIN 304
   | :||| | :||:| | :||| | ||||| :||| | :||| |
QY 304 KHAFSGGNDTIEHRQFGNCDVVSFMYLTFPLEDDDKLEQIRKDYSGAMLTGELKKA 363
   | :||| | :||:| | :||| | ||||| :||| | :||| |
Db 305 RYAFSSGGDSIEKHRELQANLEVIDPVKYLSFLEDDDELHIKKEYGEGRLTGEVKKR 364
   | :||| | :||:| | :||| | ||||| :||| | :||| |
QY 364 LIEVLOPLIAEHOARKREVTDEIYKEFMTPKRISFDQ 401
   | ||| :| :| :| | ||||| | | | | |
Db 365 LIEVLTETIVEKRRARAATVDEMADAFMAVRPLPSKFE 402
   | ||| :| :| :| | ||||| | | | | |

RESULT 10
AAG23697
ID AAG23697 standard; Protein: 426 AA.
XX
AC AAG23697;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SMO ID NO: 27100.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 9905-0121825.
PR
```

```
PR 05-MAR-1999; 9905-0123180.
PR 09-MAR-1999; 9905-0123548.
PR 23-MAR-1999; 9905-0125788.
PR 25-MAR-1999; 9905-0126264.
PR 29-MAR-1999; 9905-0126785.
PR 01-APR-1999; 9905-0127462.
PR 06-APR-1999; 9905-0128234.
PR 08-APR-1999; 9905-0128714.
PR 16-APR-1999; 9905-0129845.
PR 19-APR-1999; 9905-0130077.
PR 21-APR-1999; 9905-0130449.
PR 23-APR-1999; 9905-0130510.
PR 28-APR-1999; 9905-0130891.
PR 30-APR-1999; 9905-0131449.
PR 30-APR-1999; 9905-0132048.
PR 30-APR-1999; 9905-0132407.
PR 04-MAY-1999; 9905-0132484.
PR 05-MAY-1999; 9905-0132485.
PR 06-MAY-1999; 9905-0132486.
PR 06-MAY-1999; 9905-0132487.
PR 07-MAY-1999; 9905-0132863.
PR 11-MAY-1999; 9905-0134256.
PR 14-MAY-1999; 9905-0134218.
PR 14-MAY-1999; 9905-0134219.
PR 14-MAY-1999; 9905-0134221.
PR 14-MAY-1999; 9905-0134370.
PR 18-MAY-1999; 9905-0134376.
PR 19-MAY-1999; 9905-0134941.
PR 20-MAY-1999; 9905-0135124.
PR 21-MAY-1999; 9905-0135353.
PR 24-MAY-1999; 9905-0135629.
PR 25-MAY-1999; 9905-0136021.
PR 27-MAY-1999; 9905-0136392.
PR 28-MAY-1999; 9905-0136782.
PR 01-JUN-1999; 9905-0137222.
PR 03-JUN-1999; 9905-0137528.
PR 04-JUN-1999; 9905-0137502.
PR 07-JUN-1999; 9905-0137724.
PR 08-JUN-1999; 9905-0138094.
PR 10-JUN-1999; 9905-0138540.
PR 10-JUN-1999; 9905-0138847.
PR 14-JUN-1999; 9905-0139119.
PR 16-JUN-1999; 9905-0139452.
PR 16-JUN-1999; 9905-0139453.
PR 17-JUN-1999; 9905-0139492.
PR 18-JUN-1999; 9905-0139454.
PR 18-JUN-1999; 9905-0139455.
PR 18-JUN-1999; 9905-0139456.
PR 18-JUN-1999; 9905-0139457.
PR 18-JUN-1999; 9905-0139458.
PR 18-JUN-1999; 9905-0139459.
PR 18-JUN-1999; 9905-0139460.
PR 18-JUN-1999; 9905-0139461.
PR 18-JUN-1999; 9905-0139462.
PR 18-JUN-1999; 9905-0139463.
PR 18-JUN-1999; 9905-0139750.
PR 18-JUN-1999; 9905-0139763.
PR 21-JUN-1999; 9905-0139817.
PR 22-JUN-1999; 9905-0139899.
PR 23-JUN-1999; 9905-0140353.
PR 23-JUN-1999; 9905-0140354.
PR 24-JUN-1999; 9905-0140695.
PR 28-JUN-1999; 9905-0140823.
PR 29-JUN-1999; 9905-0140991.
PR 30-JUN-1999; 9905-0141287.
PR 01-JUL-1999; 9905-0141842.
PR 01-JUL-1999; 9905-0142154.
PR 02-JUL-1999; 9905-0142055.
PR 06-JUL-1999; 9905-0142390.
PR 08-JUL-1999; 9905-0142803.
PR 09-JUL-1999; 9905-0142920.
PR 12-JUL-1999; 9905-0142977.
PR 13-JUL-1999; 9905-0143542.
PR
```

PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 17-JUN-1999; 99US-0139453.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139753.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142380.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 16-AUG-1999; 99US-0148684.  
PR 17-AUG-1999; 99US-0149368.  
PR 18-AUG-1999; 99US-01494175.  
PR 20-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0149950.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.





KW vascular endothelial cell function; burn; plastic surgery; abdomen;  
 KW polymorphonuclear leucocyte elastase; angiogenesis; tumor metastasis;  
 KW angiogenesis; graft; myocardial infarction; solid tumour; wound healing;  
 KW dermal ulcer; diabetic ulcer; endothelialization;  
 KW lypophanyl-tRNA synthetase; tRPS; vascular graft surgery.  
 XX Homo sapiens.  
 OS  
 PN WO200174841-A1.  
 PD 11-OCT-2001.  
 XX  
 XX 21-MAR-2001; 2001WO-US08966.  
 XX  
 XX 31-MAR-2000; 2000US-193471P.  
 XX  
 XX (SCRI ) SCRIPPS RES INST.  
 XX  
 XX Schimmel P, Wakasugi K;  
 PI  
 XX  
 DR WPI: 2001-626377/72.  
 DR N-PSDB: AAA43605.  
 XX  
 PS Disclosure: Page 135-36; 150pp; English.  
 XX  
 XX The sequences given in ABA47615-18 show full length and truncated  
 CC versions of lypophanyl-tRNA synthetase (tRPS). The truncated tRPS of  
 CC the invention comprises a Rossmann fold nucleotide binding domain, and  
 CC is capable of regulating vascular endothelial cell function. It is of  
 CC approx. 40 kilo Dalton molecular weight and is produced by cleavage of  
 CC full length tRPS with polymorphonuclear leucocyte elastase. Truncated  
 CC tRPS is useful for regulating angiogenesis, tumor metastasis,  
 CC enhancing angiogenesis to a graft, treating myocardial infarction,  
 CC solid tumor, and a condition that would benefit from increased or  
 CC decreased angiogenesis in a mammal, in particular humans. It is also  
 CC useful in diagnosis and as a wound healing agent for treating wounds  
 CC such as dermal ulcers, diabetic ulcers, burns and injuries and in  
 CC plastic surgery when reconstruction is required following a burn or  
 CC for cosmetic purposes. It is particularly useful in the treatment of  
 CC abdominal wounds where there is high risk of infection. Truncated tRPS  
 CC promotes endothelialization in vascular graft surgery and is used in  
 CC conjunction with angiography to administer the angiogenic tRNA  
 CC synthetase polypeptides or polynucleotides directly to the lumen and  
 CC wall of the blood vessel.  
 CC  
 XX  
 SQ Sequence 392 AA;  
 Query Match 94.0%; Score 1988; DB 22; Length 392;  
 Best Local Similarity 100.0%; Pred. No. 2,1e-196;  
 Matches 376; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 24 SAKGIDYDKLVFRGSSKIDKELNRIERATGQRPHFRGIFSSHDMQVLDAYENK 83  
 DB 2 SAKGIDYDKLVFRGSSKIDKELNRIERATGQRPHFRGIFSSHDMQVLDAYENK 61  
 QY 84 KPFYLYTGRGSSSEAMHVGHLIPFTFKWLDVENVPLVIOMTDEKYLMDLTLDQAYG 143  
 DB 62 KPFYLYTGRGSSSEAMHVGHLIPFTFKWLDVENVPLVIOMTDEKYLMDLTLDQAYG 121  
 QY 144 DAVENAKDIACGFDINKTFFSDDLIDYGMSSGFYKNNVKIQKHVTFNQVKGIFGFTDSD 203  
 DB 122 DAVENAKDIACGFDINKTFFSDDLIDYGMSSGFYKNNVKIQKHVTFNQVKGIFGFTDSD 181  
 QY 204 CIGKISPAIOAAPSFSFPOIFRDRDIOCLIPCAIDODPYFRMTDVAIPRIGYPPA 263  
 DB 182 CIGKISPAIOAAPSFSFPOIFRDRDIOCLIPCAIDODPYFRMTDVAIPRIGYPPA 241  
 QY 264 LLSHTFFPALOGAOTKMSADPNSSIFLTDPAKOIKTVKNHAFSGGRTIEEHROFCGN 323

DB 242 LLSHTFFPALOGAOTKMSADPNSSIFLTDPAKOIKTVKNHAFSGGRTIEEHROFCGN 301  
 QY 324 CDVDSFMYLTFLEDDDKLEOIRKDYSGAMLTGELKKALIEVLOPIAEHOARRKEVT 383  
 DB 302 CDVDSFMYLTFLEDDDKLEOIRKDYSGAMLTGELKKALIEVLOPIAEHOARRKEVT 361  
 QY 384 DEIVKEFWTPRKLSEDFQ 401  
 DB 362 DEIVKEFWTPRKLSEDFQ 379  
 RESULT 7  
 ID ABB64621 standard; Protein; 430 AA.  
 XX  
 AC ABB64621;  
 XX  
 XX 26-MAR-2002 (first entry)  
 DT  
 XX  
 DE Drosophila melanogaster polypeptide SEQ ID NO 20655.  
 XX  
 KW Drosophila; developmental biology; cell signalling; insecticide;  
 KW pharmaceutical.  
 XX  
 OS Drosophila melanogaster.  
 XX  
 PN WO200171042-A2.  
 XX  
 XX 27-SEP-2001.  
 PD  
 XX  
 XX 23-MAR-2001; 2001WO-US09231.  
 XX  
 XX 23-MAR-2000; 2000US-191637P.  
 PR 11-JUL-2000; 2000US-0614150.  
 XX  
 XX (PEKE ) PE CORP NY.  
 XX  
 PI Venter JC, Adams M, Li PMD, Myers EW;  
 DR WPI: 2001-656860/75.  
 DR N-PSDB: ABL08724.  
 XX  
 XX New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signalling and cell-cell  
 PT interactions -  
 XX  
 PS Disclosure: SEQ ID NO 20655; 21pp + Sequence Listing; English.  
 XX  
 XX The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
 CC sequences (ABL01840-ABL16175) and the encoded proteins  
 CC (ABB57737-ABB72072).  
 CC  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pcl\_sequences.  
 CC  
 XX  
 SQ Sequence 430 AA;  
 Query Match 64.8%; Score 1370.5; DB 22; Length 430;  
 Best Local Similarity 64.3%; Pred. No. 1.4e-132;  
 Matches 253; Conservative 62; Mismatches 76; Indels 1; Gaps 1;  
 QY 7 ATAEEDPVDPMVTOTSSAKGIDYDKLVFRGSSKIDKELNRIERATGQRPHFRGIF 66  
 DB 38 ATAEEDPVDPMVTOTSSAKGIDYDKLVFRGSSKIDKELNRIERATGQRPHFRGIF 97  
 QY 67 FFSDHDMQVLDAYENKPFYLYTGRGSSSEAMHVGHLIPFTFKWLDVENVPLVIOMT 126  
 DB 98 FFSDHDMQVLDAYENKPFYLYTGRGSSSEAMHVGHLIPFTFKWLDVENVPLVIOMT 157

CC associated proteins represented in AAB58106 - AAB58548. Lung cancer  
CC associated proteins and polynucleotide sequences, their agonists, and  
CC antagonists may have neuroprotective; cytostatic; cardioprotective;  
CC immunomodulatory; muscular active general; vulnery; gastrointestinal  
CC general; nephrotropic; anti-infective; gynecological; or antibacterial  
CC activity. The invention also includes antibodies specific for the  
CC protein or polynucleotide sequences. The lung cancer associated  
CC polynucleotide sequences may be used for detection of lung cancer,  
CC chromosome identification, as chromosome markers, and for numerous other  
CC diagnostic or research purposes. The proteins may be used to treat  
CC disorders such as neural, immune, muscular, reproductive,  
CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative  
CC disorders. The proteins may also be used in the treatment of wounds and  
CC infectious diseases. Polynucleotide sequences AAF18425 - AAF18433 and  
CC peptide AAB58549 are used in the course of the invention for the  
CC identification and characterisation of the polynucleotide and protein  
CC sequences.  
XX  
SQ Sequence 475 AA:  
Query Match 99.3%; Score 2101; DB 21: Length 475;  
Best Local Similarity 99.5%; Pred. No. 5.9e-208;  
Matches 399; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 SNHGPDATAEEDFVDPMTVQTSSAKGIDYDKLIYRFGSSKIDKELINRIERATGQRPHH 60  
DB 75 SNHGPDATAEEDFVDPMTVQTSSAKGIDYDKLIYRFGSSKIDKELINRIERATGQRPHH 134  
QY 61 FLRRGIFFSHRMNOVLDAVENKKPFYLYTGRGPSSSEAMHGHILPFTFKWLODVFNVP 120  
DB 135 FLRRGIFFSHRMNOVLDAVENKKPFYLYTGRGPSSSEAMHGHILPFTFKWLODVFNVP 194  
QY 121 LVIQMTDEKYLKMDLTLDQAYGDAVENAKDIIACGFDINKTFIFSDLDYMGSSGFYKN 180  
DB 135 LVIQMTDEKYLKMDLTLDQAYGDAVENAKDIIACGFDINKTFIFSDLDYMGSSGFYKN 254  
QY 181 VYIKQHTVFNQVKGIFGFTSDCIGKISFPALQAPSPNSFPQIFRDRDIOCLIPCA 240  
DB 255 VYIKQHTVFNQVKGIFGFTSDCIGKISFPALQAPSPNSFPQIFRDRDIOCLIPCA 314  
QY 241 IDDDYFPMTRDVAPRIGPKPALHSTFFPALQAGQTKMSADPNSIFLDTAKQIKT 300  
DB 315 IDDDYFPMTRDVAPRIGPKPALHSTFFPALQAGQTKMSADPNSIFLDTAKQIKT 374  
QY 301 KVNKAFFSGGRDTIEHROFGNCDVVSFMYLTFLEDDDKLEQIRKDYTSGAMLTGEL 360  
DB 375 KVNKAFFSGGRDTIEHROFGNCDVVSFMYLTFLEDDDKLEQIRKDYTSGAMLTGEL 434  
QY 361 KKALEVLQPLIAEHQARRKEVTDEIVKEFMPRKLSEDFQ 401  
DB 435 KKALEVLQPLIAEHQARRKEVTDEIVKEFMPRKLSEDFQ 475  
RESULT 5  
AAY05372  
ID AAY05372 standard: Protein: 471 AA.  
XX  
AC AAY05372:  
XX  
DT 30-JUN-1999 (first entry)  
XX  
DE Human HCMV inducible gene protein, SEQ ID NO 12.  
XX  
XX HCMV inducible gene; c1g; human; human cytomegalovirus; interferon;  
KW anti-viral therapy; anti-HCMV therapy; detection; diagnosis;  
XX drug screening.  
XX  
OS Homo sapiens.  
XX  
XX WO9913075-A2.  
XX  
XX PD 18-MAR-1999.  
XX

PF 08-SEP-1998; 98WO-US18638.  
XX  
XX 22-SEP-1997; 97US-0059725.  
PR 08-SEP-1997; 97US-0058180.  
XX  
XX (UYPR-) UNIV PRINCETON.  
XX  
PI Cong J, Schenk T, Zhu H;  
XX  
XX WPI: 1999-243729/20.  
DR N-PSDB; AAX33942.  
XX  
XX New isolated human genes  
PS Claim 3; Page 112-114; 184pp; English.  
XX  
XX This sequence is encoded by a human gene of the invention, and is induced  
CC to express by both HCMV and Interferon (IFN), designated HCMV-Inducible  
CC genes (c1g or c1gs). The invention also relates to genes that are  
CC repressed in the presence of HCMV infection, designated HCMV-repressible  
CC genes (crg or crgs). The products can be used to obtain agents which can  
CC be used for anti-viral therapy, particularly anti-HCMV therapy. They can  
CC also be used for the development of drugs that would allow for higher  
CC dosage IFN treatments without the concomitant toxicity normally  
CC associated with administering high levels of IFN. The products can also  
CC be used for detection, diagnosis and drug screening.  
XX  
SQ Sequence 471 AA:  
Query Match 99.1%; Score 2096; DB 20: Length 471;  
Best Local Similarity 99.3%; Pred. No. 1.9e-207;  
Matches 398; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 1 SNHGPDATAEEDFVDPMTVQTSSAKGIDYDKLIYRFGSSKIDKELINRIERATGQRPHH 60  
DB 71 SNHGPDATAEEDFVDPMTVQTSSAKGIDYDKLIYRFGSSKIDKELINRIERATGQRPHH 130  
QY 61 FLRRGIFFSHRMNOVLDAVENKKPFYLYTGRGPSSSEAMHGHILPFTFKWLODVFNVP 120  
DB 131 FLRRGIFFSHRMNOVLDAVENKKPFYLYTGRGPSSSEAMHGHILPFTFKWLODVFNVP 190  
QY 121 LVIQMTDEKYLKMDLTLDQAYGDAVENAKDIIACGFDINKTFIFSDLDYMGSSGFYKN 180  
DB 191 LVIQMTDEKYLKMDLTLDQAYGDAVENAKDIIACGFDINKTFIFSDLDYMGSSGFYKN 250  
QY 181 VYIKQHTVFNQVKGIFGFTSDCIGKISFPALQAPSPNSFPQIFRDRDIOCLIPCA 240  
DB 251 VYIKQHTVFNQVKGIFGFTSDCIGKISFPALQAPSPNSFPQIFRDRDIOCLIPCA 310  
QY 241 IDDDYFPMTRDVAPRIGPKPALHSTFFPALQAGQTKMSADPNSIFLDTAKQIKT 300  
DB 311 IDDDYFPMTRDVAPRIGPKPALHSTFFPALQAGQTKMSADPNSIFLDTAKQIKT 370  
QY 301 KVNKAFFSGGRDTIEHROFGNCDVVSFMYLTFLEDDDKLEQIRKDYTSGAMLTGEL 360  
DB 371 KVNKAFFSGGRDTIEHROFGNCDVVSFMYLTFLEDDDKLEQIRKDYTSGAMLTGEL 430  
QY 361 KKALEVLQPLIAEHQARRKEVTDEIVKEFMPRKLSEDFQ 401  
DB 431 KKALEVLQPLIAEHQARRKEVTDEIVKEFMPRKLSEDFQ 471  
RESULT 6  
AAB47618  
ID AAB47618 standard: Protein: 392 AA.  
XX  
AC AAB47618:  
XX  
DT 07-JAN-2002 (first entry)  
XX  
XX Human inactive TTPRS.  
XX  
XX Tyrosyl-tRNA synthetase; TYRS; Rossmann fold nucleotide binding domain;

QY 301 KVNKHAFIGSGRDTIEHROFGNCDVVSFMYLTFELEDDEKLEIRKDYTSGAMLTGEL 360  
DB 324 KVNKHAFIGSGRDTIEHROFGNCDVVSFMYLTFELEDDEKLEIRKDYTSGAMLTGEL 383  
QY 361 KKALIEVLQPLIAEHQARRKEVTDEIVKEFMTPRKLSDFQ 401  
DB 384 KKALIEVLQPLIAEHQARRKEVTDEIVKEFMTPRKLSDFQ 424

RESULT 3  
ID AAB47615 standard; Protein: 484 AA.  
AC AAB47615:  
XX 07-JAN-2002 (first entry)  
DE Human full-length TrpRS.  
XX Tyrosyl-tRNA synthetase: TyRS: Rossmann fold nucleotide binding domain;  
KW vascular endothelial cell function; burn; plastic surgery; abdomen;  
KW polymorphonuclear leucocyte elastase; angiogenesis; tumour metastasis;  
KW angiogenesis; graft; myocardial infarction; solid tumour; wound healing;  
KW dermal ulcer; diabetic ulcer; endothelialization;  
KW tryptophanyl-tRNA synthetase; trpRS; vascular graft surgery.  
XX Homo sapiens.  
XX MO200174841-A1.  
PD 11-OCT-2001.  
PE 21-MAR-2001; 2001WO-US08966.  
XX 31-MAR-2000; 2000US-193471P.  
PR (SCRI ) SCRIPPS RES INST.  
PA Schimmel P, Wakasugi K;  
PI Schimmel P, Wakasugi K;  
XX WPI: 2001-626377/72.  
DR N-PSDB: AAF43602.  
XX New human truncated tyrosyl-tRNA synthetase polypeptide for regulating  
PT vascular endothelial function, in particular for regulating  
PT angiogenesis, tumor metastasis and treating myocardial infarction -  
XX  
PS Disclosure: Page 117-19; 150pp; English.  
XX  
CC The sequences given in AAB47615-18 show full length and truncated  
CC versions of tryptophanyl-tRNA synthetase (TrpRS). The truncated TrpRS of  
CC the invention comprises a Rossmann fold nucleotide binding domain, and  
CC is capable of regulating vascular endothelial cell function. It is of  
CC approx. 40 kDa molecular weight and is produced by cleavage of  
CC full length TrpRS with polymorphonuclear leucocyte elastase. Truncated  
CC TrpRS is useful for regulating angiogenesis, tumor metastasis,  
CC enhancing angiogenesis to a graft, treating myocardial infarction,  
CC solid tumor, and a condition that would benefit from increased or  
CC decreased angiogenesis in a mammal, in particular humans. It is also  
CC useful in diagnosis and as a wound healing agent for treating wounds  
CC such as dermal ulcers, diabetic ulcers, burns and injuries and in  
CC plastic surgery when reconstruction is required following a burn or  
CC for cosmetic purposes. It is particularly useful in the treatment of  
CC abdominal wounds where there is high risk of infection. Truncated TrpRS  
CC promotes endothelialization in vascular graft surgery and is used in  
CC conjunction with angiography to administer the angiogenic tRNA  
CC synthetase polypeptides or polynucleotides directly to the lumen and  
CC wall of the blood vessel.  
XX  
SQ Sequence 484 AA;  
Query Match 100.0%; Score 2116; DB 22; Length 484;

Best Local Similarity 100.0%; Pred. No. 1,7e-209;  
Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 SNHGDATEAEEDPVPMTVQTSKAGIDYDKLYRVSSSKIDKLIRIERATGQRRPH 60  
DB 71 SNHGDATEAEEDPVPMTVQTSKAGIDYDKLYRVSSSKIDKLIRIERATGQRRPH 130  
QY 61 FLRRCIFSSHDMNOVLDAVENKKRPFYLYTGRGPSSSEAMHGHLPFTFTKWLQDFVNP 120  
DB 131 FLRRCIFSSHDMNOVLDAVENKKRPFYLYTGRGPSSSEAMHGHLPFTFTKWLQDFVNP 190  
QY 121 LVIQMTDEKYLWKDLTLDQAYGDAVENAKDIACGPDINKTFTFSLDYMGMSGFYKN 180  
DB 191 LVIQMTDEKYLWKDLTLDQAYGDAVENAKDIACGPDINKTFTFSLDYMGMSGFYKN 250  
QY 181 VKIOKHVTENOVKICIFGFTSDICGKISFPAIQAAFSFSNFPQIFRDRDIOCLIPCA 240  
DB 251 VKIOKHVTENOVKICIFGFTSDICGKISFPAIQAAFSFSNFPQIFRDRDIOCLIPCA 310  
QY 241 IDODEYFRMTDVAAPRIGYPPALHSTFPALQAGOTKMSASDPNSSIFLTDPAKOIKT 300  
DB 311 IDODEYFRMTDVAAPRIGYPPALHSTFPALQAGOTKMSASDPNSSIFLTDPAKOIKT 370  
QY 301 KVNKHAFIGSGRDTIEHROFGNCDVVSFMYLTFELEDDEKLEIRKDYTSGAMLTGEL 360  
DB 371 KVNKHAFIGSGRDTIEHROFGNCDVVSFMYLTFELEDDEKLEIRKDYTSGAMLTGEL 430  
QY 361 KKALIEVLQPLIAEHQARRKEVTDEIVKEFMTPRKLSDFQ 401  
DB 431 KKALIEVLQPLIAEHQARRKEVTDEIVKEFMTPRKLSDFQ 471

RESULT 4  
ID AAB58220 standard; Protein: 475 AA.  
AC AAB58220:  
XX 14-MAR-2001 (first entry)  
DE Lung cancer associated polypeptide sequence SEQ ID 558.  
XX  
XX Human; lung cancer associated protein; neuroprotective; cytostatic;  
KW cardiactive; immunomodulatory; muscular active; vulnerary;  
KW gastrointestinal; nephrotropic; antiinfective; gynecological;  
KW antibacterial; diagnosis; neural disorder; immune disorder; reproductive;  
KW proliferative disorder; wound healing; infectious disease.  
XX  
OS Homo sapiens.  
XX  
PN WO200055180-A2.  
XX  
XX 21-SEP-2000.  
PD 08-MAR-2000; 2000WO-US05918.  
PE 12-MAR-1999; 99US-0124270.  
XX  
PR (HUMA-) HUMAN GENOME SCI INC.  
PA (ROSE/) ROSEN C A.  
PI Ruben SM;  
XX WPI: 2000-587514/55.  
DR N-PSDB: AAF18096.  
XX Lung cancer associated gene sequences, referred to as lung cancer  
PT antigens, useful for treatment, prevention, and diagnosis of disorders  
PT such as lung cancer -  
XX  
XX Claim 11; Page 1052-1053; 1425pp; English.  
XX Polynucleotide sequences AAF17982 - AAF18424 encode human lung cancer

PT New human truncated tyrosyl-tRNA synthetase polypeptide for regulating  
PT vascular endothelial function, in particular for regulating  
PT angiogenesis, tumor metastasis and treating myocardial infarction -  
PS  
PS Disclosure; Page 129-30; 150pp; English.  
XX  
XX The sequences given in ABA47615-18 show full length and truncated  
CC versions of triphophanyl-tRNA synthetase (TrpRS). The truncated TrpRS of  
CC the invention comprises a Rosemann fold nucleotide binding domain, and  
CC is capable of regulating vascular endothelial cell function. It is of  
CC approx. 40 kilo Dalton molecular weight and is produced by cleavage of  
CC full length TrpRS with polynucleonuclear leucocyte elastase. Truncated  
CC TrpRS is useful for regulating angiogenesis, tumor metastasis,  
CC enhancing angiogenesis to a graft, treating myocardial infarction,  
CC solid tumor, and a condition that would benefit from increased or  
CC decreased angiogenesis in a mammal, in particular humans. It is also  
CC useful in diagnosis and as a wound healing agent for treating wounds  
CC such as dermal ulcers, diabetic ulcers, burns and injuries and in  
CC plastic surgery when reconstruction is required following a burn or  
CC for cosmetic purposes. It is particularly useful in the treatment of  
CC abdominal wounds where there is high risk of infection. Truncated TrpRS  
CC promotes endothelialization in vascular graft surgery and is used in  
CC conjunction with angiography to administer the angiogenic tRNA  
CC synthetase polypeptides or polynucleotides directly to the lumen and  
CC wall of the blood vessel.  
XX  
XX Sequence 415 AA:  
SQ  
Query Match 100.0%; Score 2116; DB 22; Length 415;  
Best Local Similarity 100.0%; Pred. No. 1.4e-209;  
Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 SNHGDATEAEDEVDPTVOTSSAKGIDYDKLIYRFSSKIDKELNIRERATGQRPNH 60  
DB 2 SNHGDATEAEDEVDPTVOTSSAKGIDYDKLIYRFSSKIDKELNIRERATGQRPNH 61  
QY 61 FLRRCIFFSHRDMNOVDAYENKKRFYLYTGRGSSSEAMHVGHLPIFTKWLQDVFNVP 120  
DB 62 FLRRCIFFSHRDMNOVDAYENKKRFYLYTGRGSSSEAMHVGHLPIFTKWLQDVFNVP 121  
QY 121 LVIOQMTDEKYLKMDLTLDQAYGDAVENAKDIIACGPIINTFTFSDIDYMGSSGFYKN 180  
DB 122 LVIOQMTDEKYLKMDLTLDQAYGDAVENAKDIIACGPIINTFTFSDIDYMGSSGFYKN 181  
QY 181 VKIÖKHVTFNQVKGIFGFTSDICIGKISPAIOAAPSFSNFPQIFRDRIQCLLPCA 240  
DB 182 VKIÖKHVTFNQVKGIFGFTSDICIGKISPAIOAAPSFSNFPQIFRDRIQCLLPCA 241  
QY 241 IDÖPYFRMTDVAAPRIGYPRPALHSTFFPALOGAÖTKMSASDPNSSIFLDTAKÖIKT 300  
DB 242 IDÖPYFRMTDVAAPRIGYPRPALHSTFFPALOGAÖTKMSASDPNSSIFLDTAKÖIKT 301  
QY 301 KVNKHAFFSGGDDTIEHHQFGNCDVDSFMYLTFLEDDDKLEDIRKDYSGAMLTGEL 360  
DB 302 KVNKHAFFSGGDDTIEHHQFGNCDVDSFMYLTFLEDDDKLEDIRKDYSGAMLTGEL 361  
QY 361 KKALIEVLOPLIAEHQARKKEVTDEIVKEFMTPKLSFDQ 401  
DB 362 KKALIEVLOPLIAEHQARKKEVTDEIVKEFMTPKLSFDQ 402  
RESULT 2  
ABA47616 ID ABA47616 standard; Protein; 437 AA.  
XX  
XX ABA47616;  
XX  
XX 07-JAN-2002 (first entry)  
XX  
XX Human mini TrpRS.  
XX  
XX Tyrosyl-tRNA synthetase; TyRS; Rosemann fold nucleotide binding domain;  
KW vascular endothelial cell function; burn; plastic surgery; abdomen;

KW polymorphonuclear leucocyte elastase; angiogenesis; tumor metastasis;  
KW angiogenesis; graft; myocardial infarction; solid tumour; wound healing;  
KW dermal ulcer; diabetic ulcer; endothelialization;  
KW triphophanyl-tRNA synthetase; trpRS; vascular graft surgery.  
XX  
XX Homo sapiens.  
OS  
PN WO200174841-A1.  
PN  
XX 11-Oct-2001.  
PD  
XX 21-MAR-2001; 2001WO-US08966.  
PF  
XX 31-MAR-2000; 2000US-193471P.  
PR  
XX (SCRI ) SCRIPPS RES INST.  
PA  
XX Schimmel P, Wakasugi K;  
XX  
XX WPI; 2001-626377/72.  
DR  
XX N-PSDB; AAH43603.  
PT  
PT New human truncated tyrosyl-tRNA synthetase polypeptide for regulating  
PT vascular endothelial function, in particular for regulating  
PT angiogenesis, tumor metastasis and treating myocardial infarction -  
PS  
PS Disclosure; Page 123-24; 150pp; English.  
XX  
XX The sequences given in ABA47615-18 show full length and truncated  
CC versions of triphophanyl-tRNA synthetase (TrpRS). The truncated TrpRS of  
CC the invention comprises a Rosemann fold nucleotide binding domain, and  
CC is capable of regulating vascular endothelial cell function. It is of  
CC approx. 40 kilo Dalton molecular weight and is produced by cleavage of  
CC full length TrpRS with polynucleonuclear leucocyte elastase. Truncated  
CC TrpRS is useful for regulating angiogenesis, tumor metastasis,  
CC enhancing angiogenesis to a graft, treating myocardial infarction,  
CC solid tumor, and a condition that would benefit from increased or  
CC decreased angiogenesis in a mammal, in particular humans. It is also  
CC useful in diagnosis and as a wound healing agent for treating wounds  
CC such as dermal ulcers, diabetic ulcers, burns and injuries and in  
CC plastic surgery when reconstruction is required following a burn or  
CC for cosmetic purposes. It is particularly useful in the treatment of  
CC abdominal wounds where there is high risk of infection. Truncated TrpRS  
CC promotes endothelialization in vascular graft surgery and is used in  
CC conjunction with angiography to administer the angiogenic tRNA  
CC synthetase polypeptides or polynucleotides directly to the lumen and  
CC wall of the blood vessel.  
XX  
XX Sequence 437 AA;  
SQ  
Query Match 100.0%; Score 2116; DB 22; Length 437;  
Best Local Similarity 100.0%; Pred. No. 1.5e-209;  
Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 SNHGDATEAEDEVDPTVOTSSAKGIDYDKLIYRFSSKIDKELNIRERATGQRPNH 60  
DB 24 SNHGDATEAEDEVDPTVOTSSAKGIDYDKLIYRFSSKIDKELNIRERATGQRPNH 83  
QY 61 FLRRCIFFSHRDMNOVDAYENKKRFYLYTGRGSSSEAMHVGHLPIFTKWLQDVFNVP 120  
DB 84 FLRRCIFFSHRDMNOVDAYENKKRFYLYTGRGSSSEAMHVGHLPIFTKWLQDVFNVP 143  
QY 121 LVIOQMTDEKYLKMDLTLDQAYGDAVENAKDIIACGPIINTFTFSDIDYMGSSGFYKN 180  
DB 144 LVIOQMTDEKYLKMDLTLDQAYGDAVENAKDIIACGPIINTFTFSDIDYMGSSGFYKN 203  
QY 181 VKIÖKHVTFNQVKGIFGFTSDICIGKISPAIOAAPSFSNFPQIFRDRIQCLLPCA 240  
DB 204 VKIÖKHVTFNQVKGIFGFTSDICIGKISPAIOAAPSFSNFPQIFRDRIQCLLPCA 263  
QY 241 IDÖPYFRMTDVAAPRIGYPRPALHSTFFPALOGAÖTKMSASDPNSSIFLDTAKÖIKT 300  
DB 264 IDÖPYFRMTDVAAPRIGYPRPALHSTFFPALOGAÖTKMSASDPNSSIFLDTAKÖIKT 323

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 24, 2002, 12:49:40 ; Search time 25.1199 Seconds  
(without alignments)  
1773.117 Million cell updates/sec

Title: US-09-813-718-10\_COPY\_71\_471  
Perfect score: 2116  
Sequence: 1 SNHGPDATAEDEFDVPMTV.....VTDEIVKEMTPKLSFDQ 401

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: /SIDSI/gcgdata/geneseq/emb1/AA1980.DAT:\*  
2: /SIDSI/gcgdata/geneseq/emb1/AA1981.DAT:\*  
3: /SIDSI/gcgdata/geneseq/emb1/AA1982.DAT:\*  
4: /SIDSI/gcgdata/geneseq/emb1/AA1983.DAT:\*  
5: /SIDSI/gcgdata/geneseq/emb1/AA1984.DAT:\*  
6: /SIDSI/gcgdata/geneseq/emb1/AA1985.DAT:\*  
7: /SIDSI/gcgdata/geneseq/emb1/AA1986.DAT:\*  
8: /SIDSI/gcgdata/geneseq/emb1/AA1987.DAT:\*  
9: /SIDSI/gcgdata/geneseq/emb1/AA1988.DAT:\*  
10: /SIDSI/gcgdata/geneseq/emb1/AA1989.DAT:\*  
11: /SIDSI/gcgdata/geneseq/emb1/AA1990.DAT:\*  
12: /SIDSI/gcgdata/geneseq/emb1/AA1991.DAT:\*  
13: /SIDSI/gcgdata/geneseq/emb1/AA1992.DAT:\*  
14: /SIDSI/gcgdata/geneseq/emb1/AA1993.DAT:\*  
15: /SIDSI/gcgdata/geneseq/emb1/AA1994.DAT:\*  
16: /SIDSI/gcgdata/geneseq/emb1/AA1995.DAT:\*  
17: /SIDSI/gcgdata/geneseq/emb1/AA1996.DAT:\*  
18: /SIDSI/gcgdata/geneseq/emb1/AA1997.DAT:\*  
19: /SIDSI/gcgdata/geneseq/emb1/AA1998.DAT:\*  
20: /SIDSI/gcgdata/geneseq/emb1/AA1999.DAT:\*  
21: /SIDSI/gcgdata/geneseq/emb1/AA2000.DAT:\*  
22: /SIDSI/gcgdata/geneseq/emb1/AA2001.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2116	100.0	415	22	AA847617 Human supermini tr
2	2116	100.0	437	22	AA847616 Human mini trPRS.
3	2116	100.0	484	22	AA847615 Human full-length
4	2101	99.3	475	21	AA858220 Lung cancer associ
5	2096	99.1	471	20	AA858219 Human HCMV inducib
6	1988	94.0	392	22	AA847618 Human inactive Trp
7	1370.5	64.8	430	22	AB864621 Drosophila melanog
8	1370.5	64.8	430	22	AB867203 Drosophila melanog
9	1304	61.6	402	21	AA823698 Arabidopsis thalia
10	1304	61.6	426	21	AA823697 Arabidopsis thalia
11	1218.5	57.6	424	22	AA866931 Tryptophanyl-LRNA

12	968	45.7	292	21	AA823699 Arabidopsis thalia
13	831	39.3	173	21	AA858517 Lung cancer associ
14	803	37.9	385	22	AA896409 putative P. abyssi
15	226.5	10.7	142	21	AA858219 Lung cancer associ
16	185.5	8.8	341	19	AA856423 Tryptophanyl-LRNA
17	185.5	8.8	341	22	AA837851 Streptococcus pneu
18	164.5	7.8	409	20	AA894248 C. albicans tyrosy
19	162.5	7.7	409	20	AA894247 C. albicans tyrosy
20	159.5	7.5	379	20	AA837623 protein involved i
21	158.5	7.5	378	22	AA896600 putative P. abyssi
22	152	7.2	344	20	AA835439 Chlamydia pneumoni
23	146	6.9	339	22	AA835829 Helicobacter pylor
24	138.5	6.5	140	21	AA823032 Arabidopsis thalia
25	134.5	6.4	525	22	AA860745 Drosophila melanog
26	134	6.3	372	22	AA847614 Human mini trPRS m
27	132.5	6.3	372	22	AA847612 Human mini trPRS.
28	132.5	6.3	356	22	AA847611 Human full length
29	128	6.0	334	22	AA835477 Haemophilus influe
30	123	5.8	347	21	AA831904 Arabidopsis thalia
31	122.5	5.8	346	21	AA848617 Arabidopsis thalia
32	118.5	5.6	404	22	AA854020 Propionibacterium
33	115.5	5.5	129	21	AA823033 Arabidopsis thalia
34	113.5	5.4	418	22	AA820578 Enterococcus faeca
35	113.5	5.4	423	22	AA835085 Enterococcus faeca
36	113	5.3	415	22	AA833998 Staphylococcus aur
37	113	5.3	420	18	AA819781 Tyrosyl-LRNA synth
38	113	5.3	420	22	AA837056 Staphylococcus aur
39	112	5.3	345	22	AA830498 C. glutamicum prote
40	108	5.1	334	22	AA838213 Salmonella typhi c
41	106.5	5.0	334	22	AA834778 E. coli cellular p
42	104.5	4.9	370	20	AA817508 Pneumocystis carin
43	100.5	4.7	421	22	AA882363 S. epidermidis ope
44	95	4.5	418	18	AA827663 Streptococcus pneu
45	95	4.5	418	22	AA837869 Streptococcus pneu

## ALIGNMENTS

RESULT 1	AA847617 standard; Protein: 415 AA.
XX	AA847617:
AC	07-JAN-2002 (first entry)
XX	
DE	Human supermini trPRS.
XX	
KW	Tyrosyl-LRNA synthetase; TyRS; Rossmann fold nucleotide binding domain;
KW	vascular endothelial cell function; burn; plastic surgery; abdomen;
KW	polymorphonuclear leucocyte elastase; angiogenesis; tumour metastasis;
KW	angiogenesis; graft; myocardial infarction; solid tumour; wound healing;
KW	dermal ulcer; diabetic ulcer; endothelialization;
KW	tryptophanyl-LRNA synthetase; trPRS; vascular graft surgery.
XX	
OS	Homo sapiens.
PN	WO200174841-A1.
XX	
PD	11-OCT-2001.
XX	
PF	21-MAR-2001; 2001WO-US08966.
XX	
PR	31-MAR-2000; 2000US-193471P.
XX	
PA	(SCRI ) SCRIPPS RES INST.
XX	
PI	Schimmel P, Wakasugi K;
XX	
DR	WPI: 2001-626377/72.
XX	
DR	N-PSDB: AA843604.
XX	

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 24, 2002, 12:49:45 ; Search time 14.2521 Seconds  
(without alignments)  
2858.658 Million cell updates/sec

Title: US-09-813-718-10\_COPY\_48\_471

Perfect score: 2246  
Sequence: 1 MSKRAAGEDTKADCPGPNP.....VTDEIVKEFMTPKLSFDQ 424

Scoring table: BLOSUM62  
Gapop 10.0 , Gapept 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR-71:\*

- 1: PIR1:\*
- 2: PIR2:\*
- 3: PIR3:\*
- 4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2231	99.3	471	1 A41706	tryptophan--trna 1
2	2134.5	95.0	475	1 YWBO	tryptophan--trna 1
3	2041	90.9	475	1 YWRBPR	tryptophan--trna 1
4	2032	90.5	481	2 S50053	tryptophan--trna 1
5	1210	53.9	395	2 S58157	hypothetical prote
6	1163	51.8	432	2 S51901	tryptophan--trna 1
7	907	40.4	386	2 C90190	tryptophan--trna 1
8	803	35.8	385	2 C75020	tryptophan--trna 1
9	626.5	27.9	301	2 G71206	tryptophan--trna 1
10	534.5	23.8	380	2 G84373	tryptophan--trna 1
11	409.5	18.2	370	2 P64476	tryptophan--trna 1
12	397.5	17.7	364	2 E69131	tryptophan--trna 1
13	386	17.2	134	2 T43806	tryptophan--trna 1
14	370.5	16.5	420	2 E69461	tryptophan--trna 1
15	353	15.7	374	2 D72477	probable tryptopha
16	274.5	12.2	513	2 F84371	tryptophan--trna
17	192	8.5	323	2 H69346	tryptophan--trna
18	188	8.4	364	2 E72512	probable tyrosyl-t
19	185.5	8.3	341	2 D95250	tryptophan--trna 1
20	185.5	8.3	341	2 G98155	tryptophan--trna 1
21	174.5	7.8	341	2 B86633	tryptophan--trna 1
22	172	7.7	394	2 A45999	tyrosine--trna 119
23	170.5	7.6	366	2 S75410	tyrosine--trna 119
24	169	7.5	395	2 H70385	tryptophan--trna 1
25	167	7.4	351	2 E75438	tryptophan--trna 1
26	166.5	7.4	346	2 B71496	tryptophan--trna 1
27	164.5	7.3	895	2 A86410	protein F3M18.22
28	163.5	7.3	353	2 E70100	tryptophan--trna 1
29	163	7.3	460	2 C64750	probable tyrosyl-t

30	159.5	7.1	337	2 F71300	tryptophan--trna 1
31	158.5	7.1	375	2 B75072	tyrosyl--trna synth
32	158	7.0	346	2 C81654	tryptophan--trna
33	154.5	6.9	102	2 T44994	probable tryptopha
34	152	6.8	344	2 H86590	tryptophan--trna
35	152	6.8	344	2 C72034	tryptophan--trna 1
36	149.5	6.7	327	2 C84374	tyrosyl--trna synth
37	149.5	6.7	408	2 T03741	probable tyrosine-
38	146	6.5	339	2 E64676	tryptophan--trna 1
39	146	6.5	375	2 F71093	tyrosine--trna 119
40	145	6.5	337	2 A11066	tryptophan--trna 1
41	138	6.1	319	2 H69102	tyrosine--trna 119
42	135.5	6.0	328	2 C72370	tryptophan--trna 1
43	135	6.0	365	2 E82052	tryptophan--trna 1
44	132.5	5.9	343	2 S73024	tryptophan--trna 1
45	132	5.9	328	1 YWBSF	tryptophan--trna 1

ALIGNMENTS

RESULT 1  
A41706  
tryptophan--trna ligase (EC 6.1.1.2) [similarity] - human  
N:Alternate names: Interferon-inducible protein IFP53; peptide-chain release factor 1  
C:Species: Homo sapiens (man)  
C>Date: 19-May-2000 #sequence\_revision 19-May-2000 #ext\_change 19-May-2000  
C:Accession: A41633; A41706; S19246; JN0676; JH0533; S26287  
R:Flückner, J.; Rasmussen, H.H.; Justesen, J.  
Proc. Natl. Acad. Sci. U.S.A. 88, 11520-11524, 1991  
A:Title: Human Interferon gamma potentially induces the synthesis of a 55-kDa protein (C  
A:Reference number: A41633; M01D:92107982  
A:Accession: A41633  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-471 <RUB>  
A:Cross-references: GB:X59892; NID:930820; PIDN:CAA42545.1; PID:930821  
R:Rubin, B.Y.; Anderson, S.L.; Xing, L.; Powell, R.J.; Tate, W.P.  
J. Biol. Chem. 266, 24245-24248, 1991  
A:Title: Interferon induces tryptophanyl-trna synthetase expression in human fibrobl  
A:Reference number: A41706; M01D:92105071  
A:Accession: A41706  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-471 <RUB>  
A:Cross-references: EMBL:X62570; NID:932708; PIDN:CAA44450.1; PID:932709  
R:Prolova, L.Y.; Grigorjeva, A.Y.; Sudomolina, M.A.; Kisselev, L.L.  
Gene 188, 237-246, 1993  
A:Title: The human gene encoding tryptophanyl-trna synthetase: Interferon-response e  
A:Reference number: JN0676; M01D:93292932  
A:Accession: JN0676  
A:Molecule type: DNA  
A:Residues: 1-141:182-471 <FRO1>  
A:Cross-references: GB:X67918; GB:S62837; NID:937968; GB:X67919; NID:937969; GB:X679  
4; NID:937974; GB:X67925; GB:S62855; NID:937975; GB:X67926; GB:S62856; NID:937976; GE  
A:Note: the authors translated the codon GGG for residue 55 as Cys and GAG for residu  
R:Prolova, L.Y.; Sudomolina, M.A.; Grigorjeva, A.Y.; Zinovleva, O.L.; Kisselev, L.L.  
Gene 109, 291-296, 1991  
A:Title: Cloning and nucleotide sequence of the structural gene encoding for human tr  
A:Reference number: JH0533; M01D:92112058  
A:Accession: JH0533  
A:Molecule type: mRNA





A:Residues: 1-212, 'GD', 215-471 <FR02>  
 A:Cross-references: GB:M61715; NID:9340367; PIDN:AAA61298.1; PID:9340368  
 A:Experimental source: fibroblast  
 C:Genetics:  
 A:Gene: GDB:MARS; IFP53  
 A:Cross-references: GDB:119632; OMTM:191050  
 A:Map position: 14q23-14q31  
 A:Introns: 33/3; 105/1; 141/2; 181/2; 242/2; 276/1; 313/3; 371/3; 418/3  
 C:Superfamily: mammalian tryptophan--tRNA ligase; amino acid--tRNA ligase repeat homolog  
 C:Keywords: aminocyl-tRNA synthetase; ATP; ligase; protein biosynthesis  
 F:19-64/Domain: amino acid--tRNA ligase repeat homology <ATL>

Query Match 99.3%; Score 2231; DB 1; Length 471;  
 Best Local Similarity 99.5%; Pred. No. 1,4e-173;  
 Matches 422; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY 1 MSYKAAGEDYKADCPGNPAPTNSHGPDATEAEEDFVDPMTVOTSSAKGIDYKLYRG 60
DB 48 MSYKAAGEDYKADCPGNPAPTNSHGPDATEAEEDFVDPMTVOTSSAKGIDYKLYRG 107
61 GSSKIDKELINRIERATGQRPHFLRGIFFSHRDMNOVLAYENKKPFYLYTGRGSSA 120
DB 108 GSSKIDKELINRIERATGQRPHFLRGIFFSHRDMNOVLAYENKKPFYLYTGRGSSA 167
QY 121 AMYGHILPIFTFKLQDVNVPVLYOMTDDEKYLKDLDOAYGDAVENAKDIACGF 180
DB 168 AMYGHILPIFTFKLQDVNVPVLYOMTDDEKYLKDLDOAYGDAVENAKDIACGF 227
QY 181 DINKTFISDLIDYMGSSGFYKNVYKIOKHVTFNOVYGFISFSDSGIKISFPALQAP 240
DB 228 DINKTFISDLIDYMGSSGFYKNVYKIOKHVTFNOVYGFISFSDSGIKISFPALQAP 287
QY 241 SFSNSFPOIFRDRTDIOCLIPCAIDODPYFRMTRDVAPRIGYPRPALHSTFPALQAG 300
DB 288 SFSNSFPOIFRDRTDIOCLIPCAIDODPYFRMTRDVAPRIGYPRPALHSTFPALQAG 347
QY 301 TKMSADPNSISIFLTPAKOIKTKVKNHAFSGGRDTEENHOGFNCDDVDSMYLTFFLE 360
DB 348 TKMSADPNSISIFLTPAKOIKTKVKNHAFSGGRDTEENHOGFNCDDVDSMYLTFFLE 407
QY 361 EDDDKLEQIRKDYTSGLMTGELKKALIEVLOPLIAEHQARRKEVTEIYKEEPTPKLS 420
DB 408 EDDDKLEQIRKDYTSGLMTGELKKALIEVLOPLIAEHQARRKEVTEIYKEEPTPKLS 467
QY 421 FDFQ 424
DB 468 FDFQ 471

```

## RESULT 2

tryptophan--tRNA ligase (EC 6.1.1.2) [validated] - bovine  
 N:Alternate names: tryptophanyl-tRNA synthetase  
 C:Species: Bos primigenius taurus (cattle)  
 C:Date: 30-Sep-1992 #sequence, revision 30-Sep-1992 #text, change 21-Jul-2000  
 C:Accession: A40279; JN0354; S10460; S14540  
 R:Garret, M.; Pajot, B.; Trezeuet, V.; Labouesse, J.; Merle, M.; Gandar, J.C.; Benedetti  
 Biochemistry 30, 7809-7817, 1991  
 A:Title: A mammalian tryptophanyl-tRNA synthetase shows little homology to prokaryotic  
 A:Reference number: A40279; MUID:91329348  
 A:Accession: A40279  
 A:Molecule type: mRNA  
 A:Residues: 1-475 <GAR>  
 A:Cross-references: GB:M74074; EMBL:X53918; NID:9163798; PIDN:AAA30799.1; PID:9163799  
 A:Experimental source: pancreas  
 A:Note: the authors translated the codon CTG for residue 347 as Ala and CAG for residue  
 A:Note: part of this sequence was confirmed by protein sequencing  
 R:Zargava, T.A.; Kovaleva, G.K.; Pavlova, O.O.; Aevina, H.B.; Telezhinskaya, I.H.  
 Biorg. Khim. 15, 1307-1311, 1989  
 A:Title: Amino acid sequence of several peptides of tryptophanyl-tRNA synthetase from ca  
 A:Reference number: JN0354; MUID:90211408  
 A:Accession: JN0354  
 A:Molecule type: protein

A:Residues: 112-124;282-287, 'N', 288, 'F', 289-292, 'Q', 293-294, 'IR', 336-353;423-441,443-  
 A:Experimental source: liver  
 A:Note: this paper is in Russian  
 C:Superfamily: mammalian tryptophan--tRNA ligase; amino acid--tRNA ligase repeat homo  
 C:Keywords: aminocyl-tRNA synthetase; ATP; ligase; protein biosynthesis  
 F:24-69/Domain: amino acid--tRNA ligase repeat homology <ATL>

Query Match 95.0%; Score 2134.5; DB 1; Length 475;  
 Best Local Similarity 95.0%; Pred. No. 1e-165;  
 Matches 402; Conservative 9; Mismatches 11; Indels 1; Gaps 1;

```

QY 2 SYKAAGEDYKADCPGNPAPTNSHGPDATEAEEDFVDPMTVOTSSAKGIDYKLYRG 61
DB 54 SYKAAGEDYKADCPGNPAPTNSHGPDATEAEEDFVDPMTVOTSSAKGIDYKLYRG 113
62 SSKIDKELINRIERATGQRPHFLRGIFFSHRDMNOVLAYENKKPFYLYTGRGSSA 121
DB 114 SSKIDKELINRIERATGQRPHFLRGIFFSHRDMNOVLAYENKKPFYLYTGRGSSA 173
QY 122 MYGHILPIFTFKLQDVNVPVLYOMTDDEKYLKDLDOAYGDAVENAKDIACGF 181
DB 174 MYGHILPIFTFKLQDVNVPVLYOMTDDEKYLKDLDOAYGDAVENAKDIACGF 232
QY 182 INKTFISDLIDYMGSSGFYKNVYKIOKHVTFNOVYGFISFSDSGIKISFPALQAP 241
DB 233 INKTFISDLIDYMGSSGFYKNVYKIOKHVTFNOVYGFISFSDSGIKISFPALQAP 292
QY 242 SFSNSFPOIFRDRTDIOCLIPCAIDODPYFRMTRDVAPRIGYPRPALHSTFPALQAG 301
DB 293 SFSNSFPOIFRDRTDIOCLIPCAIDODPYFRMTRDVAPRIGYPRPALHSTFPALQAG 352
QY 302 KMSADPNSISIFLTPAKOIKTKVKNHAFSGGRDTEENHOGFNCDDVDSMYLTFFLE 361
DB 353 KMSADPNSISIFLTPAKOIKTKVKNHAFSGGRDTEENHOGFNCDDVDSMYLTFFLE 412
QY 362 DDDKLEQIRKDYTSGLMTGELKKALIEVLOPLIAEHQARRKEVTEIYKEEPTPKLS 421
DB 413 DDDKLEQIRKDYTSGLMTGELKKALIEVLOPLIAEHQARRKEVTEIYKEEPTPKLS 472
QY 422 DFG 424
DB 473 DFG 475

```

## RESULT 3

tryptophan--tRNA ligase (EC 6.1.1.2) [validated] - rabbit  
 N:Alternate names: tryptophanyl-tRNA synthetase  
 C:Species: Oryctolagus cuniculus (domestic rabbit)  
 C:Date: 30-Sep-1992 #sequence, revision 13-Feb-1998 #text, change 26-May-2000  
 C:Accession: A35904; S37396  
 R:Lee, C.C.; Craig, W.T.; Muzny, D.M.; Harlow, E.; Caskey, C.T.  
 Proc. Natl. Acad. Sci. U.S.A. 87, 3508-3512, 1990  
 A:Title: Cloning and expression of a mammalian peptide chain release factor with sequ  
 A:Reference number: A35904; MUID:90239043  
 A:Accession: A35904  
 A:Molecule type: mRNA  
 A:Residues: 1-475 <LEE>  
 A:Cross-references: GB:M33460  
 R:Polova, L.Y.; Dalphin, M.E.; Justesen, J.; Powell, R.J.; Druegon, G.; McCaughan, K  
 EMBO J. 12, 4013-4019, 1993  
 A:Title: Mammalian polypeptide chain release factor and tryptophanyl-tRNA synthetase  
 A:Reference number: S37396; MUID:94009008  
 A:Accession: S37396  
 A:Molecule type: mRNA  
 A:Residues: 166-177 <PRO>  
 C:Genetics:  
 A:Gene: WRS  
 C:Complex: homodimer [validated, MUID:94009008]  
 C:Function:  
 A:Description: EC 6.1.1.2 [validated, MUID:94009008]; catalyzes the ATP-dependent for  
 A:Note: mammalian WRS (tryptophanyl-tRNA synthetase) and eRF (polypeptide chain rele  
 C:Superfamily: mammalian tryptophan--tRNA ligase; amino acid--tRNA ligase repeat homo

